Effects of Internal and External Factors on Driver’s Mental Workload

March ２０１７

AHMAD KHUSHAIRY BIN MAKHTAR

Effects of Internal and External Factors on Driver's Mental Workload

Graduate School of Systems and Information Engineering
University of Tsukuba

March　２０１７

AHMAD KHUSHAIRY BIN MAKHTAR

Abstract

Although many researchers have studied the characteristics of mental workload for a long time, there is still a lot of unknown regarding mental workload. This study aims to explore the factors in increasing driver's mental workload with consideration of internal and external factors as well other factors such as driving experience. To achieve the aims of the study, three experiments were designed and conducted.

Experiment#1 focused on how to increase the accuracy of workload estimation, by (1) Normalizing data, and (2) cross-evaluation between Physiological, Performance and Subjective measurements. We were able to improve the workload estimation accuracy by normalizing the physiological measurement data. Through the cross-evaluation, we were also able to distinguish between the normal driving condition and driving under the influences of internal and external factors of mental workload.

Experiment#2 was designed to focus on determining the level of difficulties in internal factors in increasing mental workload. Two levels of difficulty in the secondary task were distinguished to represent an internal factor in mental workload, and two levels of complexity of traffic conditions as an external factor in mental workload. The results showed that the main effect of the secondary task was found on physiological, performance and subjective measurements of mental workload. We are also certain that the more complex the content of mental task is the higher workload can be seen on the workload measurements.

In Experiment#3 we investigated the influences of external factors and driving experience on mental workload. Five traffic conditions were compared. The result of lateral control of vehicle showed that the most difficult external factors in this study were driving on the winding road. The mental workload was further increased when driving under a difficult level of the mathematical task. The results also showed that experience in driving could improve performance by performing better in lateral control of the vehicle.

Overall conclusion, through this study we are able to validate the hypothetical relationship between performance and mental workload in the perspective of factors of increasing mental workload. We also confirmed that the combination of internal and external factors would increase the mental workload. This study is useful for enhancing our understanding of the mechanism of how performance deteriorate when the mental workload is high.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Driver's Mental Workload

Driving a vehicle is a complicated task. A driver not only has to master the skill to control the vehicle, but also to monitor the surroundings, especially on road traffic. A split second of carelessness will throw drivers to meet with a fatal road traffic accident. In the year 2015 Police Bureau of agency in Japan reported that distraction while driving, including 'Looking away of road' and 'Thinking of something other than driving' was 25.1% of fatal crashes in Japan [1]. This type of road traffic crashes is strongly related to 'mental workload' management [2,3]. Nowadays the usage of in-vehicle devices such as a car phone, a navigation system, and other new technology gadgets has become a trend. These features as one of the factors in increasing mental workload will further worsen the situation. The discussing driver's mental workload is worth to be one of the key elements to reducing the number of traffic crashes.

Particularly in traffic researches, managing mental workload is an important issue to be addressed because it is related to driving performance. A simple explanation of the relationship between performance and the mental workload is when a driver is in a high mental workload condition performance will be deteriorated. In driving conditions, drivers are prone to commit errors, e.g. failure in noticing a sudden brake of a leading vehicle, and this kind of errors leads to traffic accidents. Many studies have shown that a higher workload would prolong the reaction time [4–8].

Despite an important topic to be discussed, and been studied by many researchers for a long time, our knowledge regarding mental workload is not much improved. For example, the definition of mental workload itself is still ambiguous. Therefore, our study tries to provide explanations regarding mental workload and its relation to the driving activity.

1.2 Research Background

Although there is no agreeable definition about mental workload, most of the researchers interpreted mental workload as the amount of demand while performing the task. Mental workload involved the causes and the effects. Many types of research focusing on the results of increasing workloads, could be found from physiological, performance or subjective measurements. It is also an important topic to discuss what the causes of the mental workload are.

The factors in increasing mental workload could be found from inside and outside of the driver. To further discuss the factors in increasing mental workload, we divided the factors into three: 1: Internal factors, 2: External factors, and 3: others.

Internal factors could be found inside the driver when a driver engages with thinking something other than driving-related tasks while driving. Such as communicating with in-vehicle devices, conversation with other passengers, making a call, or even not focusing on driving tasks. We still have the problem of "looked but did not see" or cognitive distraction which can be translated as although their sight is on the road; they are not aware of the situation. This cognitive distraction has found to be demanding a higher level of mental workload [9–12].

On the other hand, an external factor in workload can be found outside the driver such as traffic condition or outside environment such as the geometry of the road while driving. For example, driving down a street with a high traffic load condition may increase the driver's workload [13, 14]. Driving on the curved road is more demanding than driving on a straight road[15]. Another example is that driving on a narrow road is more demanding than on a wide road [16, 17]. These kinds of road environment caused complex driving tasks and demand a higher mental workload.

While internal and external factors involved with a short-term effect and occurred in the situation, there is another factor that affects the level of mental workload which is a long-term factor. Other factors than internal and external factors as the background of drivers, the age or experience in driving itself will also affect the workload a driver perceived. However, no research concluded these factors altogether systematically by focusing on a combination of these factors and their effect to driver's mental workload. Is the combination of these factors will further increase the mental workload and whether one factor gives more effect to driver's than another are some of the questions worth to be investigated for increasing knowledge about mental workload.

Another unexplored point about mental workload is that there has not been a standard value or precise characterization of the levels of mental workload. While researchers are still in the process of finding a suitable measurement as a robust and can be used in the real driving situation, the characterization of high mental workload has not yet been discovered precisely.

There are a lot of research studies about the effect of increasing workload factor separately. This research particularly tries to combine these factors altogether to give an insight of the situation where drivers are in high internal factors mental workload while in the difficult situation of driving task (external factors). One example of the application of the knowledge from this research is to design a support system based on where the factor of mental workload come from. For instance, when driving in high traffic load, it is useful to give warning earlier to drivers so that they can engage in the situation faster.

1.3 Problem Statement

The research is motivated by these five identified problems:

1. Although numerous investigations have been done on every mental workload measurement, the robustness of the measurement has remained notably scarce.
2. Lack of usage of Blood Volume Pulse (BVP) as a candidate to evaluate driver's mental workload measurement tool.
3. Lack of study regarding factors of increasing internal and external factors of mental workload.
4. Lack of explanation about the relationship between mental workload and performance.
5. There are no standard criteria that can be determined as 'high mental workload'.

1.4 Objectives

The **main aim** of this research is to improve understanding of the characteristics of the increasing mental workload from internal and external factors.

Specifically, the research is conducted to fulfill the following objectives:

1. To assess and analyze data of driver's mental workload by using **physiological, performance and subjective measurements**.

2. To evaluate the usage of **Blood Volume Pulse** as a tool to measure mental workload.

3. To assess the **level of difficulties** of mental workload by **normalizing data** and in a systematic way **match with the existing model of mental workload**.

4. To study the relationship between increased workload and driving performance by manipulating <u>**levels of difficulties of an internal factor**</u> in mental workload.

5. To investigate the relationship between increased workload and driving performance by manipulating <u>**levels of difficulties of an external factor**</u> in mental workload.

1.5 Research Questions

This study tried to answer the following issues:

1. How do various control parameters such traffic conditions, and task workload affect the drivers?

2. Can Blood Volume Pulse be used as a tool to evaluate the driver's mental workload in driving condition?
3. What is the best parameter for physiological measurements to indicate the drivers are in the higher mental workload level?
4. Which one from internal or external factors give more effect to increasing mental workload?
5. Will additional of internal and external factors further increase mental workload?
6. What is the relationship between performance and mental workload?

1.6 Significance of Research

This research will enhance our understanding of driver's mental workload. This study will give a clearer explanation of the relationship between driver's mental workload and driving performance. This research also will fill the gap of study about possible factors of increasing mental workload. The findings from this research will help practitioners in designing a system which can detect high mental workload from external and internal factors of mental workload. The findings of this research will also be beneficial for the road engineers in designing a road which does not impose a high mental workload to drivers.

1.7 Outline of the thesis

This thesis is structured such that in Chapter 2, we review the literature regarding mental workload, the theories, and theoretical framework were discussed.

In Chapter 3 we describe the method in the attempt to increase the accuracy of evaluation regarding mental workload levels. The first experiment is designed to evaluate the potential of normalizing the data to improve the accuracy of estimation. From the results, by combining between measurement groups, we map with the hypothetical model of mental workload into regions.

In chapter 4, based on the knowledge from Experiment#1 regarding increasing the accuracy of mental workload, we conduct Experiment#2. We focus on internal factors of mental workload. We also evaluate the usage of BVP sensor as a candidate to evaluate the mental workload levels.

In chapter 5, we conduct Experiment#3 to focus our discussion on external and other factors of mental increasing mental workload.

In chapter 6, we discuss significant findings regarding mental workload and conclude the study.

Chapter 2 Literature Review

2.1 Drivers and Mental Workload

Driving a vehicle involves many complex tasks. According to Michon [18], car driving associates with three levels, i.e., strategical, maneuvering and control levels. Strategical level involves with general plans, for example, the selection of the type of route. While maneuvering, the level has a shorter time constant in seconds and on this level, a driver maneuvered the vehicle such as controlling the vehicle for not stepping out from a cruising lane. A control level with a shorter time in milliseconds relates to automatic action such as a respond to a front vehicle braking. In all of these levels, a driver has to be in good condition to maintain the safety of the vehicle. In particular, a split second of carelessness will lead to increase the risk of meeting with an accident.

One of the important criteria of a good condition of drivers is the adequate level of mental workload. Mental workload relates to the performance of drivers in controlling the vehicle. Driving in a high mental workload is not good to drivers because it makes the performance to become worse [19][20].

2.2 Mental workload theories and models

Many studies relate the theories of the mental workload with performance. Regarding the theories of mental workload, it can be divided into two categories, single resource, and multiple resources. The dominant theory in single resource theory is the capacity theory of attention by Kahneman [21]. The main idea of this theory is that there is a pool of attention and the capacity of the pool is limited. Illustrating the available capacity in a fixed pool, miscellaneous sources of arousal such as anxiety, fear, anger, and excitement will affect the performance. There is an allocation policy to which activity the allocation should be allocated. Kahneman also argued that attention also depends on the task demand. If the task is undemanding, a human operator can perform the task concurrently. However, if the task is demanding, they have to allocate the most attention to one and perform the task in serial. This condition results in the performance to be affected. This theory explains how an operator allocates their attention based on the demand of a task. In a driving condition, as driving itself involves multiple tasks, the demand on a task will give an effect to the way drivers allocating their attentional resources. If a task is a superior demand than another, a driver has to allocate more resources, and this will affect the driving performance to deteriorate. This state also identifies as the trade-offs between the primary task and secondary task performance.

Contradictory to the single resource theory, the multiple resources theory claims that an operator has multiple attentional resources. This concept was introduced by Wickens

[22], and it describes that some tasks can be performed concurrently. The resources have been represented in three dimensions one is processing stage, such as perception, cognition and responding. The second issue is input modalities. This point is another strong point of the multiple resources, which to elaborate the input modalities will also affect the performance of the task. For example, two tasks can be performed simultaneously without showing a decrement if it's been presented in different input modalities such as visually and auditory. The third one is that separate resources represent whether a task is processed verbally or spatially. This theory presented that manual and vocal responses can be efficiently time-shared. In a driving condition, talking (verbal) on the phone may be not disturbing for driving task (manual) because it is processed by a different resource.

Malleable attentional resources theory by Young & Stanton [23] tried to explain in detail the effects of underload on performance. They argued that instead of representing the resources in a fixed pool the performance also depends on malleable attentional capacity. This argument explains that reducing the demand does not necessarily improve the performance and vice versa. Recently, Young et al. [24] published a new insight regarding mental workload by adapting the model from de Waard [20]. By putting into a frame of the upper limit of resources, they designed the relationship between activation level, workload and performance model. They also highlighted the importance of discussing the 'workload red line' or overload.

Although these theories provided us to understand the relationship between mental workload and performance, to the best of our knowledge, none of them focused on factors of mental workload. In our study, we want to enhance our understanding regarding mental workload, especially help us to understand from the perspective of the factors of increasing mental workload.

2.3 Mental Workload Factors

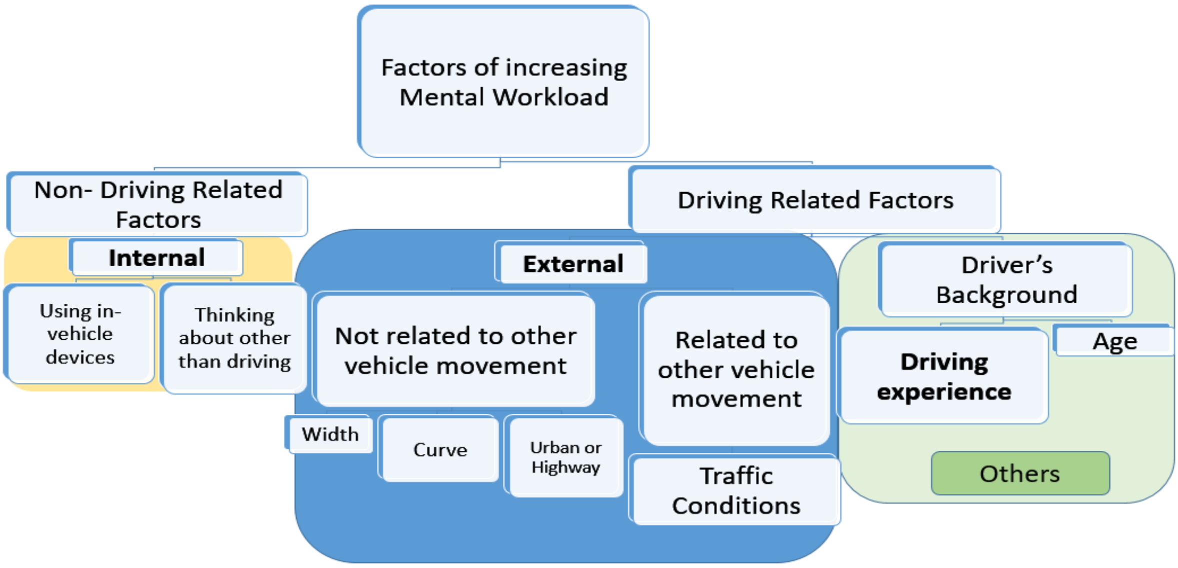


Figure 2.1 Factors of increasing mental workload

Figure 2.1 shows factors of increasing mental workload. The factors mainly can be divided into two categories, Non-driving related factors, and driving-related factors. Non –Driving related factors such as thinking about other than driving tasks or using in-vehicle devices such as mobile phone while driving.

Driving related factors can be further divided into external factors and driver's background. External factors can be divided into two, Not related to other vehicle movements such as the width of the road, curve, and the type of road either urban or highway. External factors related to other vehicle movement such as traffic conditions. The focus of our study is related to internal and external factors of increasing mental workload.

2.3.1 External Factors

In our study, external factors in increasing workload are defined as driving-related factors and can be found outside of vehicles which related to other vehicle movement such as traffic conditions or not related to other vehicle movement such as such as the geometry of the road while driving.

The followings are examples of external factors.

- Width of road
- Curvy Road
- Urban or Highway Road
- Traffic Conditions

Width of Road

Some studies explored the effects of road width on the driver's mental workload. For example, de Waard et al. [25] investigated the effect of road width on driver's performance and physiological response. An experimental road was designed by reducing the width of the road using blocks. The results showed that a narrower road made drivers less swerving and the driving speed also reduced. For a physiological response, it was also found that this kind of road condition would reduce the Heart Rate Variability (HRV). These results showed that narrower road increase driver's mental workload.

Another study was conducted by Godley et al. [26] by designing three different lane widths of 3.6 m, 3.0 m, and 2.5 m. The result of the experiment also showed a decrease in Standard Deviation of Lateral Position (SDLP). A subjective rating was investigated, and the result from NASA-TLX [27] showed the narrow road gave a higher mental workload of participants. While Rosey et al. [28] examined the influence of lane width on speed profiles and lateral positions of two types of road width, 3.0 m, and 3.5 m. The results from their study showed that reducing lane width affected the drivers to drive closer to the center of the lane. Another extensive study by Dijksterhuis et al. [29] investigated the influence of lane width of four difference width of the road of 3.00 m, 2.75 m, 2.50 m and 2.25 m. The results of the study showed that SDLP decreased about 0.01 m for each 0.25 m reduction in lane width. However, no significant differences were found in the physiological measurements.

Finally, recently a study was conducted by Brouwer et al., by manipulating lane width of 2.2 m, 2.3 m and 2.4 m[17]. Regarding the lane width, they did not find an effect of lane width on RRI and RMSSD (two kinds of HRV parameters) of physiological measurements. They argued that parameters such as skin conductor, RRI and RMSSD not sensitive to different difficulty, or mental effort conditions in driving.

The literature presented here showed a variation in effects of reducing the width of the road to the mental workload measurements. It is important to make a careful consideration of designing the road so that we can confirm the effects are from reducing the width.

Curvy Road

Under this category, many studies investigated the relationship between a curve and the speed variation of drivers such as [30–33]. We will only focus on the findings on how driving on a curvy road will give a higher mental workload to drivers compared to the straight pathway. Tsimhoni et al. [34] investigated the effects of visual demand on different three curve radii of 582 m, 291 m, and 194 m. The results showed that SDLP was increased by reducing the radius while SD of the steering wheel (SDSW) was higher when radius decrease. A study from Vollrath and Totzke [35] also found similar results of increasing SDLP compared to a straight road.

Backs et al. [15] focused on cardiac measures of driver workload with and without visual occlusion by manipulating three different radii of curves 582 m, 291 m and 194 m. They found that cardiac measures were differentially affected by curve radius. Heart rate during 582 m and 291 m curves was significantly faster than during the 194 m curve.

Ben-Bassat et al. [36] manipulated the shoulder width and curves to find the effects of driver perception and behavior. By manipulating five conditions (Straight, Shallow Right, Sharp Right, Shallow Left and Sharp Left), the results revealed that SDLP was higher during sharp left conditions.

Arien et al. [37] investigated the existence of curve as calming traffic measures and the effects on driving behavior. The significant difference was found between existing and absent of curves in SDLP.

A more recent study by Du et al. [38] investigated the effects of fatigue on driving performance under three different curve radii of 200 m, 500 m, and 800 m. The result of SDLP showed the highest value found in curve radius of 200 m. The average angle of steering wheel movement also recorded as the highest under this road geometry. The SDLP of the vehicle was decreased as the radius decreased.

In all the studies reviewed here, mental workload measurements such as SDLP and SDSW indicate a higher value when driving on a curve road compared to a straight pathway. Furthermore, the size of curve radius was also found to give an effect on mental workload measurements. The smaller the radius is, the higher mental workload could be observed.

Urban or Highway Road

Son et al. [39] investigated the impact of driving in urban and highway road on two group of driver's ages. By imposing N-back test as a secondary task, they found that cognitive workload for older drivers was higher when driving in urban section road than driving on the highway road. Liu and Lee [40] assessed the effects of the traffic situation on drivers on workload. By using instrumented vehicle and mathematical-addition test as an in-vehicle task, participants were asked to drive on urban roads and motorway roads. The results showed that no significant difference was found between the types of the road on Heart Rate (HR) and SDLP. Results of steering control showed that SDSW was larger on urban roads compared to motorway roads.

One of the explanations can be given for the question of why driving in urban roads provides a higher mental workload is because of information processing complexity. Driving in urban road demands drivers to process a lot of information while maneuvering the vehicle. There were also many studies on this matter (information processing complexity while driving in the urban road) such as [37], [41], [42].

For example, Arien et al. [37] studied the differences of workload between driving in urban areas and outside urban areas. Based on the results of Peripheral Detection Test (PDT) used in the experiment, with higher reaction time, and lower hit rate in an urban area compared to outside urban area, the workload level was higher when driving in an urban area.

Traffic Condition

Another factor that will give effects to driver's mental workload is traffic condition. For example, driving in a high-density traffic will demand a higher workload than driving in low-density traffic. In a high-density traffic, drivers have to digest a lot of information and to be aware of the movement of another vehicle such as sudden brake, signal to the right or left, and the speed of another vehicle. Strayer et al. [43] manipulated the traffic density to investigate the effect of using cell-phone during driving. Results from brake reaction times, they found that there was no significant difference between the low and high traffic density. They found a significant difference when participants are conversing with a hands-free cell phone and become worst as traffic density increased. Hao et al. [44] studied the effect of traffic density on drivers' situation awareness and mental workload. They designed two kinds of traffic density; High and Low. The result showed that with higher traffic density the subjective measurement of NASA-TLX and physiological measurement of HR of participants were increased showing that a higher mental workload for high density compared to low density.

2.3.2 Internal Factors

Internal factors of workload could be found inside drivers when the drivers are thinking something other than driving. It has often referred as a cognitive distraction. In a more real driving task, we still have the problem of "looked but did not see" or cognitive distraction. Cognitive distraction will also demand a higher level of mental workload [9–12]. Furthermore, there are also studies concluded that the more complex the conversations, the worse the impacts on drivers [45–47].

Unlike external factors of mental workload, we cannot easily differentiate levels of difficulties in internal factors. For external factors, we can change the experimental conditions to create a higher demand situation while participants are driving. A careful consideration regarding selecting a mental task should be taken so that we can make an affirmation that drivers in a high level of internal factors mental workload.

2.3.3 Other factors of mental workload

Besides internal and external factors of mental workload, there are also other factors that affect the mental workload such as the background of drivers, the age, and experience of driving. While some researchers focused on the measuring visual information behavior between experience and novice drivers such as [48 & 49], they found that experienced drivers glance less than novice drivers while performing a secondary task. They pointed out that the way of attention allocation will affect their lateral position control.

Patten et al. [50] investigated the differences between experienced and low mileage drivers under cognitive workload tasks. The results showed that the low mileage drivers had a longer Peripheral Detection Task (PDT) reaction time than the experienced. They argued that by training and experience they can automate some of driving tasks to perform better.

Yang [51] studied drivers' visual behavior under visual and auditory tasks. She divided participants into three groups: 1) Less than ten years of driving experience, 2) between 11-20 years of driving experience, and 3) more than 21 years' experience. From the results of lateral position performance, the first group (less than ten years driving experience) showed a higher value of SDLP, higher value of Steering Entropy [52] and also a higher value of Steering Rehearsal Rate (RR1) than other groups. She argued that the less driving experience group had a worse lane keeping and steering behavior both in baseline and when performing secondary tasks.

Paxion et al. [53] also divided participants into three groups of 1) Drivers age of 18 years old and had driving experience less than four months, 2) Drivers age of 21 and had driving experience of 36 months, and 3) had at least eight years of driving experience. Their results showed that, with experience, they performed better in reaction time and SDLP decreased.

All of the literature reviewed here show that the experience of driving can increase driving skills and can make driving performance better.

2.4 Mental Workload Level and Driver's performance

It is important to know the relationship between driving performance and mental workload. De Waard [20] proposed a conceptual model for describing the relationship between task performance and increase or decrease of mental workload. As shown in Figure 2.2, the model was divided into six regions. On far left and far right of the model are Region D and Region C. The human operator is inactive in Region D. It could be understood that in this region, the human operator cannot respond anymore to the changes in mental workload. In the middle of the model, Region A2 is optimum for both performance and mental workload. In this region, the human operator can easily cope with the needs of any task demands and achieve a satisfactory level of performance. In regions A1 and A3, there are two kinds of efforts, which are "state-related effort" and "task-related effort". For the state related effort, in region A1, performance remains unaffected however the state of a human operator is affected by the demand due to for example consuming alcohol or in a fatigue condition. While in region A3, the human operator is effected by "task-related effort" to maintain the performance. The performance is still at the highest level as the human operator makes an effort to maintain the condition. In region B, a human operator is unlikely to maintain anymore, and performance begins to deteriorate. The degradation of performance in region B could be interpreted as the high level of workload or overload [20].

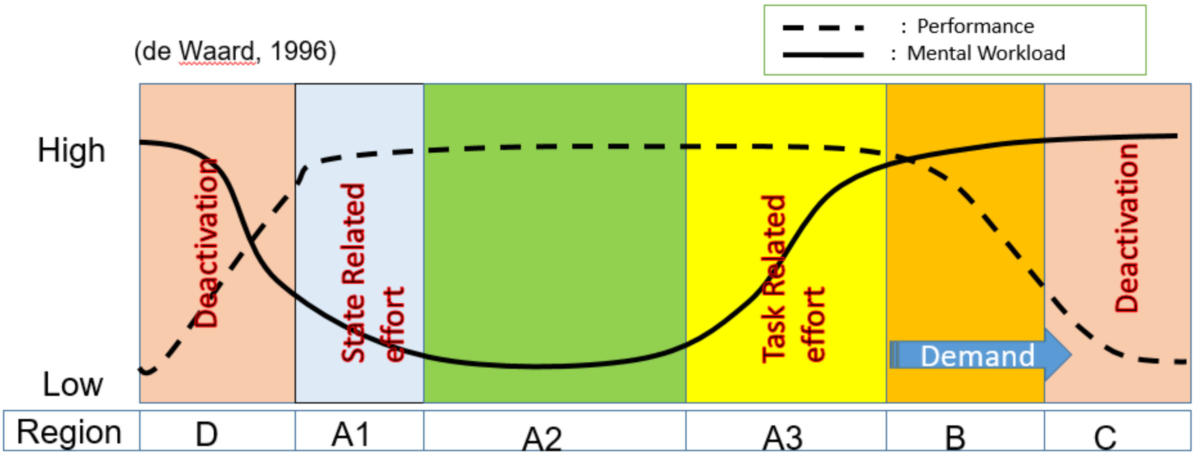


Figure 2.2 Mental Workload and Performance in 6 regions [20]

Regarding the levels of distraction, Brookhuis et al. [54] studied conversation complexity and the effects on driver distraction. Their result indicated that in simulated hands-free telephoning conversations, levels of complexity affected driver distraction. There is also a study that concluded that the more difficult and complex the conversation, the greater the possible negative effect on driver distraction [5].

The level of complexity is related with driver's cognitive distraction. In other words, the degree of task complexity imposed to the driver's cognitive distraction indicates the cognitive distraction level for drivers. It also shows that it is possible to assess the driver's cognitive distraction.

2.5 Combination of Internal and External factors in mental workload

After we have known the factors in increasing mental workload from internal and external and other factors, we come to questions such as:

1. Is there exist interaction between internal and external factors of mental workload?
2. If the interaction exists, which one gives more effect to driver's mental workload?
3. How can we improve our understanding of the relationship between mental workload and driver's performance?

To answer these questions, we draw a theoretical framework based on a hypothetical model by de Waard [20] as in Figure 2.3 which describes our hypotheses in this study towards mental workload and the performance by considering the factors of mental workload.

Based on the literature review, the factors in increasing mental workload could be found from the internal and external of drivers. The increasing level of difficulties of either the internal or external factors is expected to raise the level of mental workload further. Besides these two factors, there is also another factor that will influence the increasing mental workload such as driving experience. While Region A2 is understood as the lowest level of the mental workload while engaging with increasing demand from any factors, or the combination of these factors it is expected to increase the mental workload further.

While in Region A3, there may exist the "trade-off" between internal or external factors which one factor is more demanding than another. As discussed in de Waard's model, this situation at a certain point the operator cannot cope with the demands and the performance starts to deteriorate.

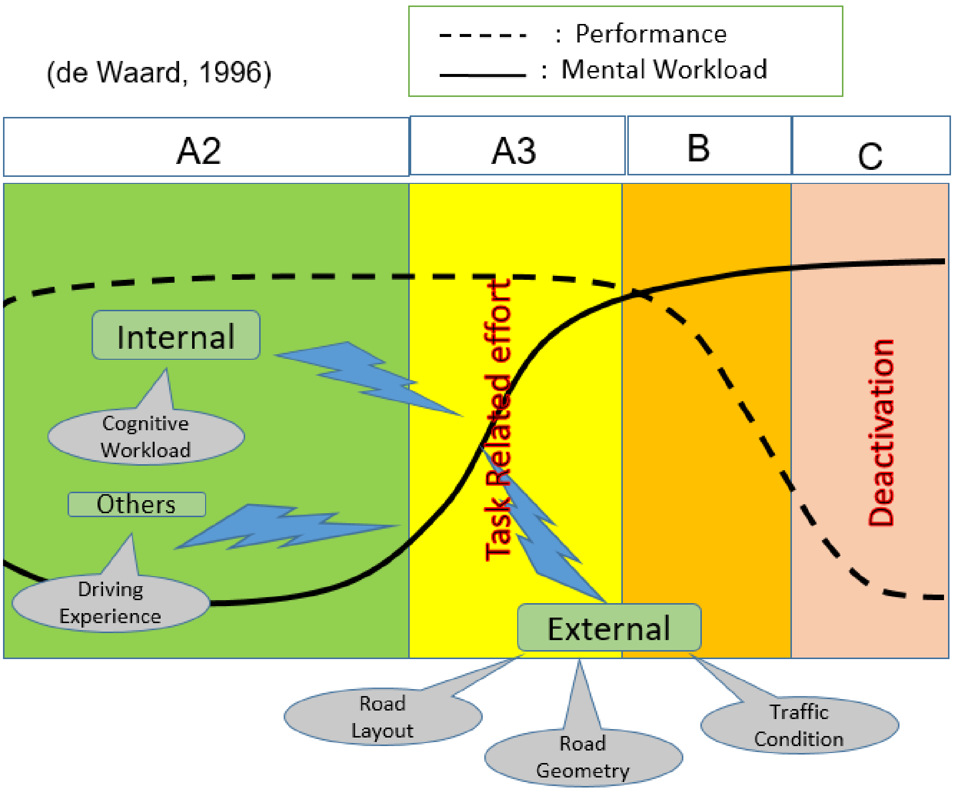


Figure 2.3 Factors of increasing mental workload considering internal and external factors

2.6 Mental Workload Measurement

Before we study further about the theoretical framework, there are still unresolved issues regarding the mental workload measurements. Based on O'Donnel and Eggemeier [55], a workload can be measured by using three measurements groups.

1. Physiological Measurements
2. Performance Measurements
3. Subjective Measurements

2.6.1 Physiological measurements

Mental workload is associated with a load on processes in the central nervous system (CNS). It contains brain, spinal cord and nervous. Any load on any part of this could give an effect to the activities of this system. As CNS in the body, it is probable that the effects could also impact another system within the body, such as respiration system, body's temperature, and other nervous systems. Monitoring the changes in any of these systems could be used as a measurement to indicate the level of mental workload. This process is the idea of how physiological measurements work. Examples of physiological measurements such as Electrocardiograph (ECG), Electromyography (EMG), Electroencephalogram (EEG), measuring blood circulation or blood pressure, etc.

The ECG is the most commonly used physiological measurements [53]. Mainly two types of the index from ECG data are said to be appropriate to measure mental workload, i.e., Heart Rate (HR) and Heart Rate Variability (HRV)[56]–[59]. Most researchers found that when opposing to higher demand of mental workload, the value of HR will increase while the value of HRV will decrease[8], [56–58], [60].

Regarding the physiological measurements, in the recent decades, mental workload's evaluation has been one of the major interesting research topics. In fact, research on mental workload estimation was carried out using numerous devices. In recent years, researches on state estimation of mental workload have been investigated with various devices such as Electroencephalogram (EEG) and Electrocardiogram (ECG) as mentioned before. However, it is quite hard to use such indices in an actual driving condition. Therefore, tools that can estimate the driver's mental workload without giving the driver a sense of restriction needed. Acknowledging the importance to develop a method to detect the driver's mental workload state, Blood Volume Pulse (BVP) sensor that records blood pulse wave is a promising candidate. BVP is non-interfere to the driving maneuvers. There is research that concluded the possibility of using a BVP sensor to quantify the degree of fatigue that results from sitting for extended periods of driving seats [61]. The usage of a BVP has been expended in driving the situation to monitor driver's blood pulse. This device located at steering of the vehicle were designed to predict heart attack while driving [62]. These facts show that a BVP sensor also can be used in a real driving situation and potential candidates for the evaluation of mental workload. Despite the potential of using a BVP sensor to estimate driver's mental workload, there are not many studies the use of blood pressure to determine workloads [63]. So, for one of the objective, our study is to evaluate the usage of BVP sensor to use in a driving situation as a measurement to evaluate driver's mental workload.

2.6.2 Performance measurements

Performance measurements refer to how well participants perform a task. A poor performance of a task indicates high mental workload which the participants struggle to perform better. There are two kinds of performance task used to measure mental workload; primary task and secondary task. In a driving situation, controlling the lateral and longitudinal position of the vehicle safe will determine how the mental workload affects the performance. Performance measurement of a primary task can be assessed by using indicators such as Standard Deviation of Lateral Position (SDLP), Standard Deviation of Steering Wheel (SDSW), Steering Entropy [52] for Lateral Position. And Time to Collision (TTC), Time Headway (THW), Reaction Time to Braking Task for Longitudinal Position.

Secondary task measures depend on the type of secondary task used to assess the mental workload. Often used secondary task performance measurement are Percentage of the Correct answer, Response time to Peripheral Detection Task, Hit rate.

2.6.3 Subjective measurements

Subjective measurement of mental workload could be performed while participants are asked to rate the workload they experienced while performing a task. While it is easy to prepare the test, it also said to be sensitive to workload except to measure the state when they were in low workload [20]. Subjective measurement techniques such as Rating Scale of Mental Effort (RSME) [64], NASA-TLX [27] and Subjective Workload Assessment Technique (SWAT) [65], have been commonly used to measure mental workload [53], [66].

2.6.4 Increasing accuracy of mental workload estimation

There are still some issues need to be improved regarding mental workload measurements. For example, until now no standard value indicates high mental workload[67]. The individual differences such as between gender and ages will make the capacity of perceiving workload is different between individual. A question on how can we validate that a driver is in high mental workload is an example of mental workload measurement issues. In this study, we are interested in increasing the accuracy of estimation regarding levels of mental workload. Two ways that we think can be implemented in this area:

1. The Cross-Evaluation between measurement groups
2. The Normalization of data

2.6.4.1 The Cross-Evaluation between measurement groups

Although there are a lot of ways to measure mental workload, the selection of measurement is crucial in a traffic research. The nature of each measurement itself which not all measurements are sensitive to all measurement groups. De Waard provided guidance on selecting measurements based on the regions [20]. For example, measuring workload redline is more appropriate to combine between measurements to get a better accuracy of estimation. Using several indicators is expected to improve the detection accuracy. Furthermore, by mapping altogether between measurements groups in one table can give us a better explanation on whether the workload is increasing or decreasing.

2.6.4.2 The Normalization of data

Another way to increase the accuracy of measurement data is by using data normalization. Because the physiological data fluctuate according to the measuring condition or physical condition of subjects, Suzuki and Okada [68] thus proposed normalizing the maximum Lyapunov exponent based on the reference value determined from the normal condition. By doing this way, we can get a standard value that somehow same to all subjects.

2.7 Summary of Literature review

Through our survey of the literature, studies on the factors in increasing mental workload is still lack. From the fact that no theories of mental workload are focusing on the factors of mental workload, we want to provide a clearer explanation about mental workload. From our revision of literature, we synthesized that the factors in mental workload could be divided into three categories: internal, external and others. We are interested to know whether the increasing difficulties in internal and external will further increase mental workload. It is also of interest to investigate the trade-off between internal and external factors on which one gives more effects to the drivers. We also want to enhance our knowledge on the relationship between mental workload and performance.

Regarding the measurements, there are some unsolved issues regarding mental workload which we think will affect the reliable of workload estimation. There are two ways we investigate to increase the accuracy of estimation. First, we have to normalize the measurement data and second the cross evaluation between measurement groups. We also want to evaluate the usage of BVP sensor as a tool to assess mental workload in a driving situation.

Chapter 3 Estimation of Driver’s Mental Workload Levels

3.1 General Overview and Aims

Although many researchers have done the study regarding mental workload, the accuracy of evaluation regarding mental workload is still notably scarce. The characterisation of high mental workload is also unclear. Until now, there is no standard value to be used as an indicator of high mental workload level. Before we discuss further mental workload, it is important for us to focus first on the method to increase the accuracy of the evaluation. As presented in the literature review, two ways that can improve the accuracy of the evaluation. (1) The cross-evaluation between measurements, and (2) The Normalization of data. We designed an experiment by focusing on these matters and taken into consideration of the level of difficulties in a combination of internal and external factors of mental workload.

The specific aims of this experiment are as follows:

- To assess and analyze data of driver's mental workload by using physiological, performance and subjective measurements.
- To distinguish levels of difficulties of the task by combining between external and internal factors in increasing mental workload.
- To evaluate the method of increasing accuracy of evaluation by a cross evaluation between groups of measurements and normalization of data

3.2 Hypotheses

The hypotheses of this experiment were:

1. The normalization of data will improve the accuracy of determination of mental workload levels by reducing the individual differences among participants.
2. The cross evaluation of measurements groups can increase the accuracy of driver’s mental workload estimation.
3. The tasks will degrade driver’s lateral control performance.

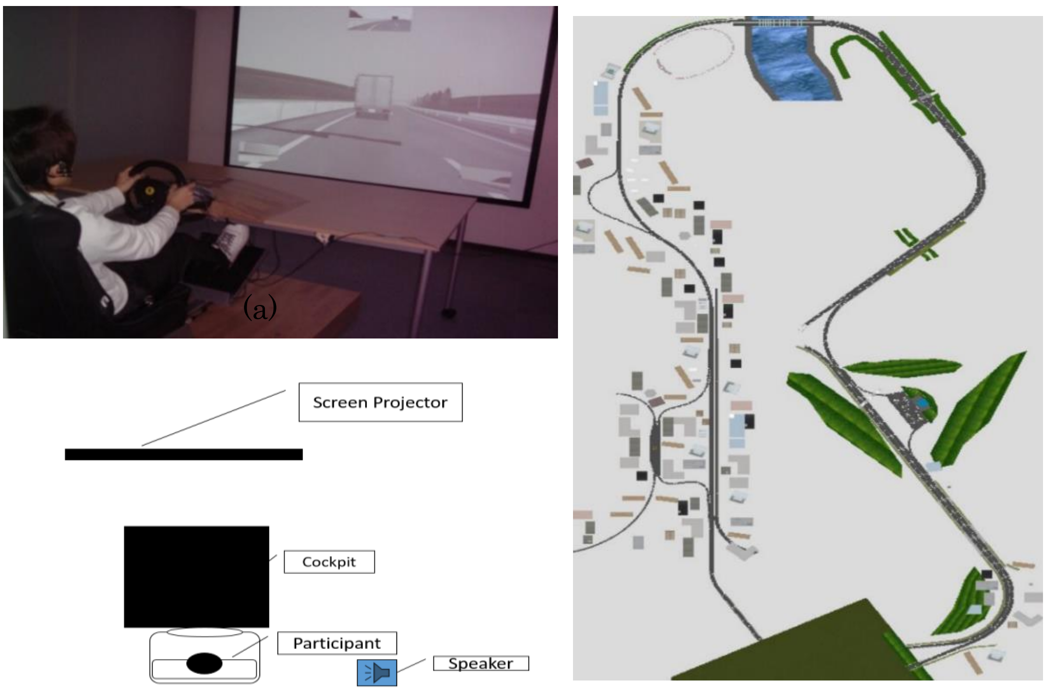
3.3 Method

3.3.1 Participants

Eight drivers (four males and four females) between the ages of 20-37 years old (Mean =23.8, SD =5.4) participated in the data collection. Every participant held a valid driver's license and drove almost daily. As this study focused on the increasing the accuracy of data, we thought that the number of participants was enough to collect the data repeatedly. With counterbalancing between genders, data of participants were sufficient enough to give a fruitful discussion regarding increasing accuracy of mental workload assessment.

3.3.2 Apparatus

The process of data collection of drivers' mental workload was done using a fixed-base driving simulator (Figure 3.1(a)(b)). Software produced by Honda Motor inc. was used to develop the scenario and parameters setting for measuring vehicle movements. For traffic environment, a highway course with 3.3 m wide lane (Figure 3.1c) was used in this experiment. The software BACS Detector was used to analyze data from two sensors of Blood Volume Pulse (BVP) were attached to ear and finger of participants.



(b) (c)

Figure 3.1 (a)(b) Driving Simulator (c) Driving course used in this experiment

3.3.3 Tasks

There were two tasks performed by the participants in this experiment.

a) The primary task as an external factor of mental workload

b) The secondary task as an internal factor of mental workload

3.3.3.1 Main Task (External factors)

As we discussed in chapter 2, one of the external factors would increase mental workload is traffic condition. In a complex situation, a driver not only has to maneuver the vehicle but also monitors the surrounding or the traffic environment including another vehicle movement. To differentiate levels of difficulties, two types of traffic complexity were designed for this experiment, namely:

a) Non-Hazardous Condition (NHC)

b) Hazardous Condition (HC).

The main task was to drive safely on the left (cruising) lane. Driving under these traffic conditions, participants were asked to drive following a Leading Vehicle (LV) (Figure 3.2). This kind of car-following task has been proved to be a reliable task to evaluate mental workload, and it appears to be more natural [8]. To control the following distance with LV, the Following Vehicle (FV) was tailgating behind the Host Vehicle (HV) driven by participants. This way, we could control the distance between LV and HV so that the participants would not open too much space between LV. There was also another vehicle cruised in the right (passing) lane.

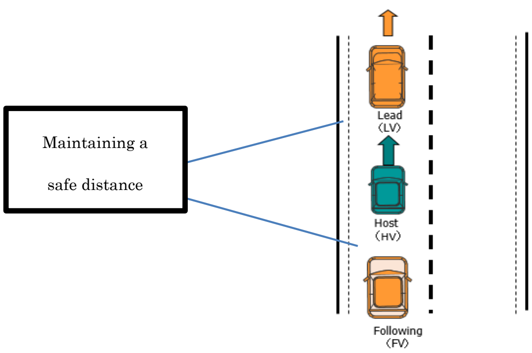


Figure 3.2 Traffic Condition

Under Non-Hazardous Condition (NHC) traffic condition, both LV and FV drove with a constant speed of 65km/h throughout 7 minutes trial.

While under HC, participants were also asked to maintain a safe distance between LV and FV. Both LV and FV cruised between speeds of 50km/h and 75km/h. Purposely, LV and FV would change the cruising speed by applying an abrupt deceleration of 0.35 G. After cruising with a lower speed for a moment, then they would accelerate again to 75km/h in all of sudden (Figure 3.3). The time point to make the speed change was set at random (approximately twice the speed changes for each 500-meter run). Overall, through seven minutes of a trial, there were twelve braking tasks made by the Lead Vehicle. Participants were asked to keep a safe following distance and at the same time to be alert with FV as it might hit from behind. If they met with a crash during a trial, participants were told to start a new trial all over again. This way, participants tried their best for not involving with a crash.

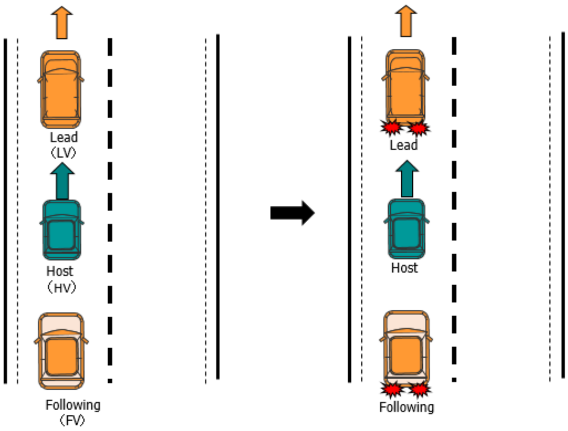


Figure 3.3 Hazardous condition (HC)

3.3.3.2 Secondary Task (Internal Factor)

We used Mathematical Arithmetic Task (MAT) to simulate as an internal factor of mental workload. MAT requires the participants to memorize the numbers presented before as well as solve the calculation. This task is a kind of so-called PASAT (Paced Auditory Serial Addition Test)[69]. Participants were requested to carry out a two-minute Mathematical Arithmetic Task (MAT) in a 7-minute trial, 3 minutes after the start and 2 minutes before the trial is completed (Figure 3.4).

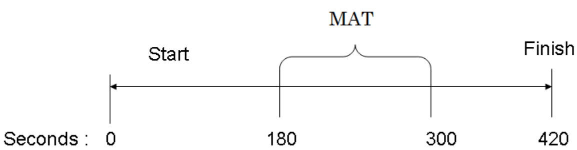


Figure 3.4 MAT period in a trial

To create a level of difficulty, MAT was divided into two levels, namely the level of easy (MAT1) and the level of difficult (MAT2) tasks.

During the two minutes of MAT period, consecutively participants hear one-digit numbers (from 1 to 9) in every three seconds through a speaker connected to a computer (see Figure 3.1(b) for the location of the speaker). With a loud voice, participants need to answer the summation of the last two numbers they heard. Overall 40 numbers were produced along the MAT period. Figure 3.5 shows an example of MAT1 (Easy level). While in MAT2, participants were given two-digit numbers between 11 and 19. Along the number of correct answers was calculated and the result was notified to participants after the end of a trial.

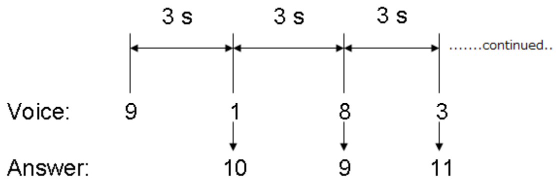


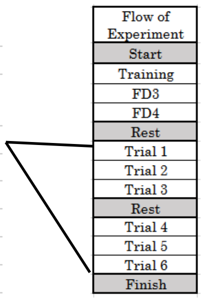
Figure 3.5 MAT1 (Easy Level)

3.3.4 Experimental Design

All participants were randomly divided into two groups (Group A and B) with 4 participants per group. It took six different days to complete the experiment. On every day of the experiment, a participant performed nine trials including training session. Table 3.1 shows the order of the experiment in one day. BD is the 'Baseline Driving', this day, participants drove in Non-Hazardous traffic condition with the absent of secondary task. FD is 'Free Driving' where there was no secondary task on the trial there was also no following vehicle exist behind the host vehicle. During FD3, participants had to follow the lead vehicle with a constant speed of 70 km/h. While in FD4, the lead vehicle will make a sudden brake and cruised at the speed of 50 km/h to 75 km/h. The purpose of designing of FD3 and FD4 was an extension to the training session. Our concern was the participants had to complete the tasks on six different days, if the separation between the days were too long, there would be differences between their driving performances. Finally, 'NMAT' is No Mathematical Arithmetic Task.

Table 3.1 Experimental procedure

Day	Group A			Group B			Number of Trials
	Secondary Task	Traffic Condition	Driving Condition	Secondary Task	Traffic Condition	Driving Condition	
One	NMAT	BD	NMATBD1	NMAT	BD	NMATBD1	**6**
Two	MAT1	NHC	MAT1NHC	MAT2	HC	MAT2HC	**6**
Three	MAT1	HC	MAT1HC	MAT2	NHC	MAT2NHC	**6**
Four	MAT2	NHC	MAT2NHC	MAT1	HC	MAT1HC	**6**
Five	MAT2	HC	MAT2HC	MAT1	NHC	MAT1NHC	**6**
Six	NMAT	BD	NMATBD2	NMAT	BD	NMATBD2	**6**



Upon arriving at the venue of the experiment, for the first day of experiment participants were explained regarding the experiment and were asked to sign a consent letter if they agreed. They were also explained about the experiment that it would take six different days including the first day and the participation was on a volunteer basis. They could withdraw from the participations whenever they want. After the explanation, they were introduced to driving simulator used in this experiment. They were given time to get familiarized with the simulator and also the secondary task with a training session. Before proceeding to the experiment, BVP device was attached to the right-hand finger of participants. As shown in Table 3.1, every participant performed six types of combination between secondary task and the traffic condition. As a subjective measurement, participants were requested to answer NASA-TLX [64] immediately after finishing every trial.

3.3.5 Dependent Variables

Table 3.2 List of Dependent Variables

Measurement Groups	Dependent Variables
Physiological	Maximum Lyapunov Exponent From Ear BVP sensor
	Maximum Lyapunov Exponent From Finger BVP sensor
Primary Task Performance	Standard Deviation of Lateral Position
Secondary task Performance	Percentage of Correct answer
Subjective	NASA-TLX

Table 3.2 shows the list of dependent variables tested in this experiment. A BVP sensor is a sensor of Blood Volume Pulse to measure changes in the volume of blood or air it contains within an organ or whole body. In this experiment, we placed the sensors to ear and finger of every participant. We placed the finger BVP sensor to the finger which hand was not dominant to participants to reduce the noise effect from the movement of the device. The maximum Lyapunov exponent was used as a parameter to reflect the changes in blood pressure in a body. This change is said to reflect the change of mental workload [67][70]. The higher the maximum Lyapunov exponential value indicates a high mental workload condition [15].

3.3.6 Data Analysis

For physiological measurement of maximum Lyapunov Exponent, three types of data analyses were performed in this experiment.

- Average Maximum Lyapunov Exponent
- Normalized Lyapunov Exponent 45s
- Normalized Lyapunov Exponent 120s

Average of Maximum Lyapunov exponent

We assumed that the demand imposed on participants during MAT interval was the same throughout two minutes of MAT interval. To quantify the participant's mental workload while performing a secondary task, we calculated the average of the Maximum Lyapunov exponent for that period.

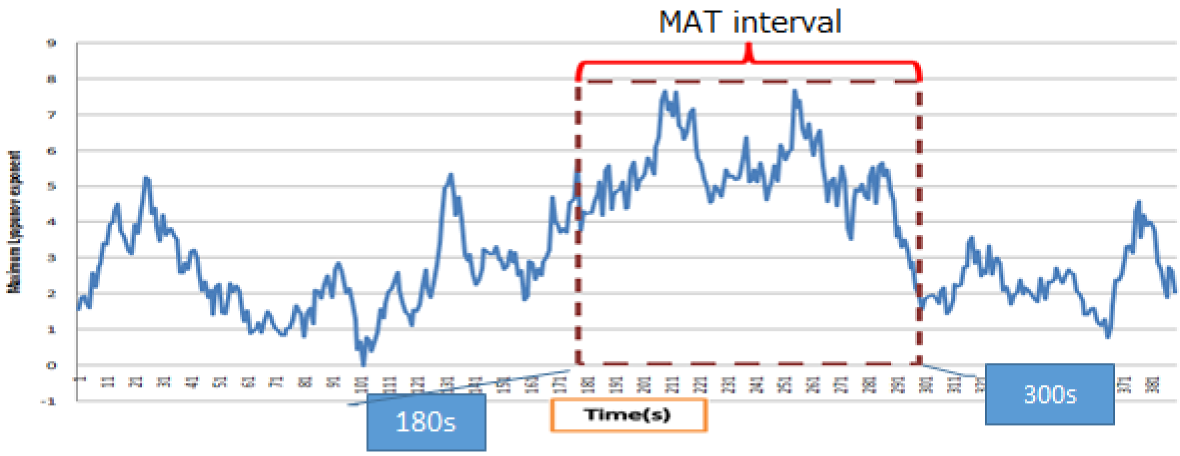


Figure 3.6 Maximum Lyapunov exponent data for a trial

Figure 3.6 showed a time-series data of the Maximum Lyapunov exponent in a trial. The time interval between 180 seconds and 300 seconds was the period when participants were performing a mathematical arithmetic task.

Normalized Maximum Lyapunov exponent

Normalized maximum Lyapunov exponent has been introduced by Suzuki and Okada [11, 13]. It is a method to increase the accuracy of mental workload evaluation by removing individual differences among participants. This method which compared data between a MAT interval data with a based period data (read: Based Interval) within the same trial performed by a particular participant.

As the method of normalizing the data itself, we think as a good way to increase the accuracy of mental workload evaluation, we still have a concern about selecting the based interval which is regarded as a reference to the MAT interval data. Our basis of concern is that the factors in increasing mental workload could be found not only from internal but also from external factors as we discussed in chapter 2. Selecting a base interval which is by the time driver is in high mental workload will, therefore, affect the accuracy of mental workload evaluation. Therefore, in this experiment, we added the information of selecting the base interval by considering two factors. 1) The length of time 2) road geometry.

In Suzuki and Okada's experiment [11, 13], they set the duration of 4 minutes within 20 minutes of driving task as a based interval. The based interval which was beginning from 3 minutes until 7 minutes from the start of experiment has been compared to MAT interval data. As there was no explanation regarding selecting the based interval been made, we considered that they made no consideration of external factors as explained above.

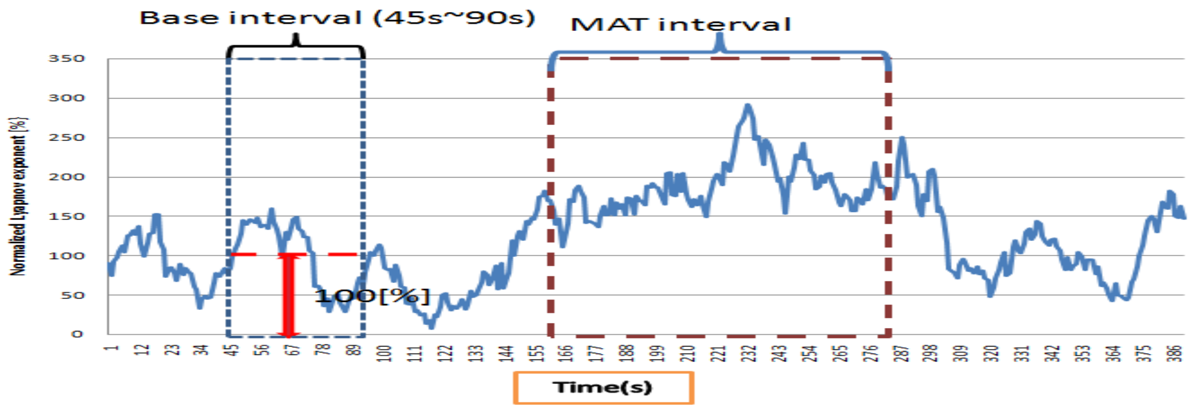
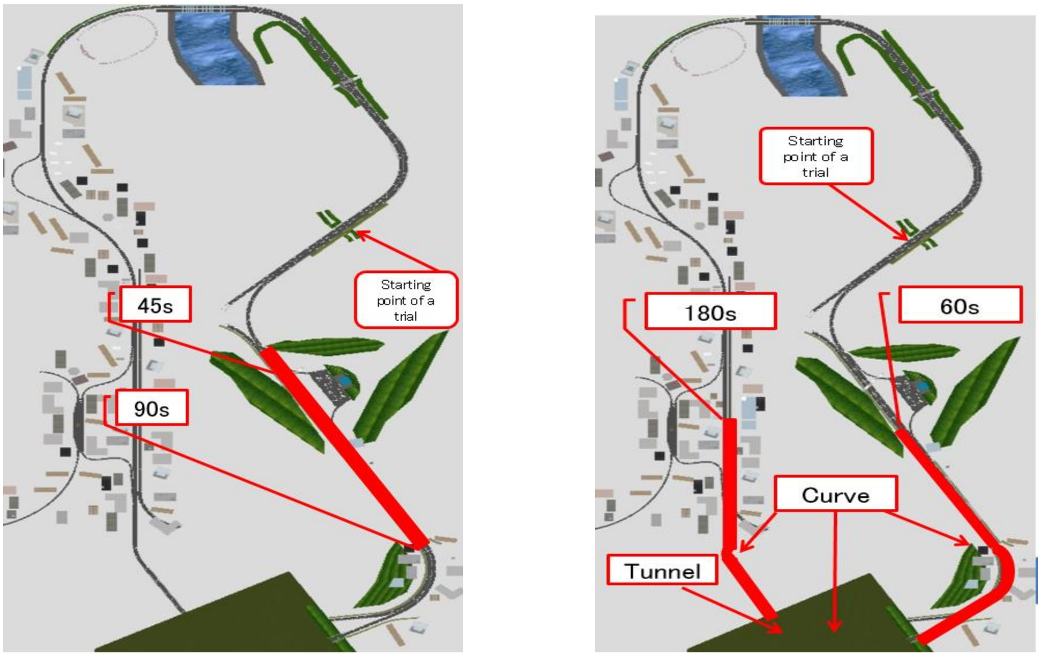


Figure 3.7 Normalized Maximum Lyapunov exponent data for one trial

Figure 3.7 shows a time series Normalized Maximum Lyapunov exponent data of a trial from one of our participant. The value of based interval was set as 100%. As mentioned before, this period will be a standard value for the MAT period of the trial. Any value higher than 100% indicates that the value was higher than the base interval.

Considering that the road geometry also will affect the mental workload measurements [9, 10], we selected the time between 45 seconds and 90 seconds of one trial as a base interval as in Figure 3.8(a). The selection was made based on the information that in this period, participants were driving in a straight pathway and no elements that can be considered as the external factors of the increasing value of the maximum Lyapunov exponent. We named this period data as "Normalized Lyapunov exponent 45s".



(a) (b)

Figure 3.8 Normalized Lyapunov Exponent

Another base interval was selected between 60 seconds 180 seconds by no consideration been made of the elements been stated before as in Figure 3.8b. This interval has also been chosen based on the same time length with the MAT interval (the interval while performing the mathematical task) (see Figure 3.9). We called this period data as "Normalized Lyapunov exponent 120s".

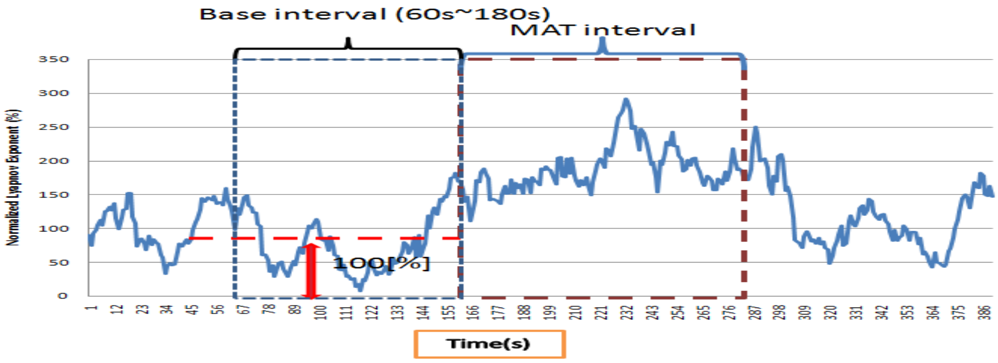


Figure 3.9 Base interval (60s~180s)

3.4 Results

3.4.1 Physiological Measurements

3.4.1.1 Maximum Lyapunov Exponent

A one-way repeated-measures ANOVA was conducted with six Driving Conditions (NMATBD1, MAT1NHC, MAT1HC, MAT2NHC, MAT2HC, and NMATBD2) as the within-subjects factor. Also, an ANOVA with A 2x2x2x8x6 mixed factorial design was done, with sources of data: Ear and Finger, Gender: Male and Female, groups: A and B, and eight participants as between-subjects factors.

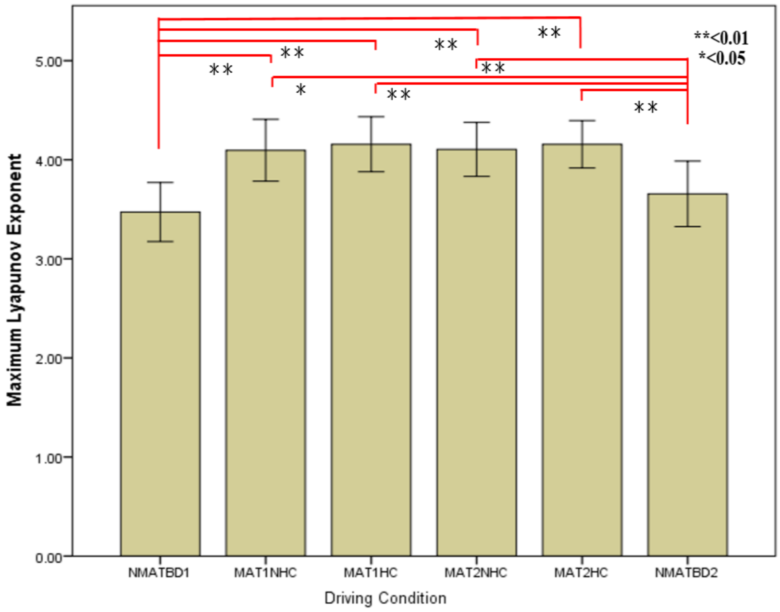


Figure 3.10 Average maximum Lyapunov exponent according to driving conditions

Figure 3.10 showed an average value of maximum Lyapunov exponent among days of an experiment for 8 participants. The results showed that there was a significant main effect of driving conditions on the value of Maximum Lyapunov exponent (F (5, 76) = 8.994, $p<0.001$, $\eta_p^2 = 0.372$). Further analysis of pairwise comparisons using Bonferroni Method revealed that significant differences were found between the driving conditions as shown in Table 3.3.

Table 3.3 Pairwise Comparisons for maximum Lyapunov Exponent. Significant effects (p<0.05) are shown in bold.

	NMATBD1	MAT1NHC	MAT1HC	MAT2NHC	MAT2HC	NMATBD2
NMATBD1	X	**.001**	**.000**	**.000**	**.000**	1.000
MAT1NHC	**.001**	X	1.000	1.000	1.000	**.029**
MAT1HC	**.000**	1.000	X	1.000	1.000	**.001**
MAT2NHC	**.000**	1.000	1.000	X	1.000	**.006**
MAT2HC	**.000**	1.000	1.000	1.000	X	**.000**
NMATBD2	1.000	**.029**	**.001**	**.006**	**.000**	X

From the results, the highest value of average Maximum Lyapunov exponent was found on the day when participants conducted the secondary task of MAT1 under HC traffic condition (M:4.156). The difference in the value of the Maximum Lyapunov exponents was statistically significant when compared with the first day of experiments, NMATBD1 (p=0.000) and between the last day of the experiment, NMATBD2 (p=0.001). Under this driving condition, participants were performing a difficult level of mathematical arithmetic task (MAT2) while driving under a complex traffic condition (HC). The value was a slightly higher than MAT2HC condition (M: 4.155).

There was also a significant difference between the first day and the day when the participants were conducting either MAT1 or MAT2 under either no hazardous condition or hazardous condition. During these days, participants were performing a mathematical arithmetic task under a complex or non-complex traffic conditions (HC or NHC).

According to the results, the values of the average maximum Lyapunov exponent of the days of MAT1NHC, MAT1HC, MAT2NHC, and MAT2HC, were quite similar and the difference between the conditions was too small. There was no significant difference between these driving conditions. From this result, it seemed that there was no effect of task difficulties level to participants.

Table 3.4 Univariate test results for maximum Lyapunov Exponent. DC = six driving conditions, Sources= Sources of data, ear or finger, Gn= Gender, G= Group A or B, P= 8 Participants. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when Mauchly's test showed a violation of sphericity.

Effect	Maximum Lyapunov Exponent			
	df 1,2	F	p	η_p^2
Driving Condition (DC)	**5,76**	**8.994**	**<0.001**	**0.372**
Sources (S)	1,80	1.711	0.282	0.014
Gender (Gn)	**1,80**	**173.608**	**<0.001**	**0.685**
Group (G)	**1,80**	**18.424**	**<0.001**	**0.187**
Participants (P)	**7,80**	**60.266**	**<0.001**	**0.841**
DC x S	**5,76**	**11.793**	**<0.001**	**0.437**
DC x P	**20,316**	**6.977**	**<0.000**	**0.306**
DC x S x P	**20,316**	**2.214**	**0.002**	**0.123**

Table 3.4 and Figure 3.11 showed the results of a univariate test of the maximum Lyapunov exponent. As shown in the table, significant effects were found on almost of the factors except for the type of sources of data. The interaction between Driving Conditions and Sources of data, between Driving conditions and participants and between the Driving conditions, Sources of data, and Participants were also found significant.

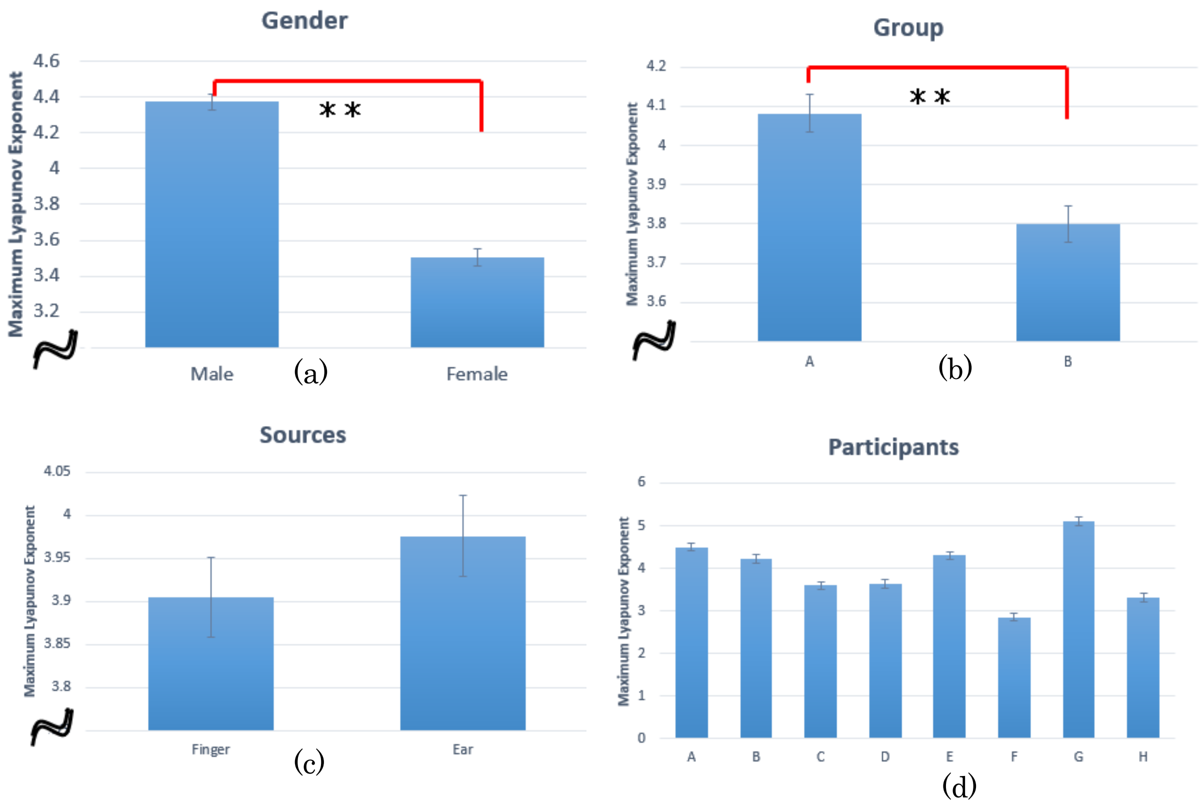


Figure 3.11 Results of univariate test of maximum Lyapunov exponent

3.4.1.2 Normalized Lyapunov Exponent

A repeated-measures ANOVA was computed from the individual means of Normalized Lyapunov Exponent in six driving conditions (NMATBD1, MAT1NHC, MAT1HC, MAT2NHC, MAT2HC, and NMATBD2). A 2x2x2x8x2x6 mixed factorial design was also performed, with the sources of data, gender, groups, participants and type of based intervals as between-subjects factors and type of driving conditions as within-subjects factors.

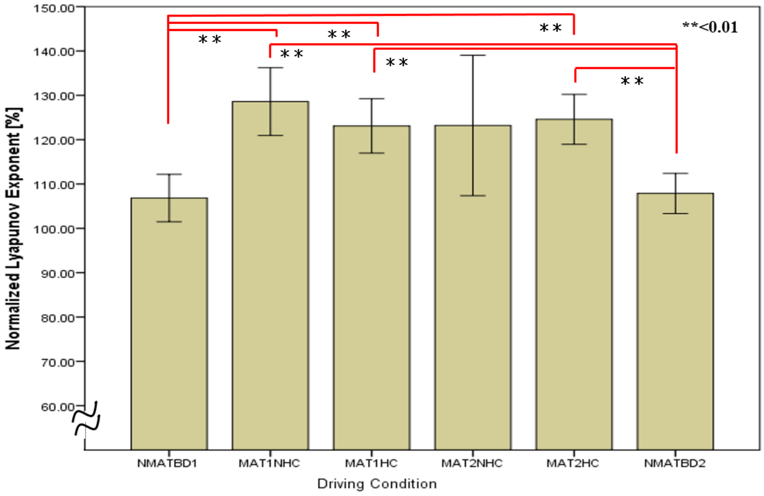


Figure 3.12 Average Normalized Lyapunov exponent based on driving conditions

Table 3.5 Pairwise comparisons for Normalized Lyapunov exponent, Significant effects (p<0.05) are shown in bold

	NMATBD1	MAT1NHC	MAT1HC	MAT2NHC	MAT2HC	NMATBD2
NMATBD1	X	**.000**	**.000**	.713	**.000**	1.000
MAT1NHC	**.000**	X	1.000	1.000	1.000	**.000**
MAT1HC	**.000**	1.000	X	1.000	1.000	**.001**
MAT2NHC	.713	1.000	1.000	X	1.000	1.000
MAT2HC	**.000**	1.000	1.000	1.000	X	**.000**
NMATBD2	1.000	**.000**	**.001**	1.000	**.000**	X

Figure 3.12 and Table 3.5 showed an average value of Normalized Lyapunov exponent and pairwise comparison for Normalized Lyapunov exponent based on driving conditions for 8 participants. The highest value of Normalized Lyapunov exponent was found on MAT1NHC driving condition (128.59%) where participants were accomplishing a one-digits number of the mathematical arithmetic task under Non-Hazardous conditions. These results were somehow different with our expectation. The second highest value

found on MAT2HC (124.582), the most difficult task in this experiment. However, as the difference between these two driving conditions was not significant, we cannot confirm that normalized Lyapunov exponent under MAT1NHC was higher than MAT2HC. There was also no significant difference between driving conditions where participants were accomplishing the secondary tasks.

Table 3.6 Univariate test results for Normalized Lyapunov Exponent. DC = six driving conditions, Sources= Sources of data, ear or finger, Gn= Gender, G= Group A or B, P= 8 Participants and Based interval: 45s or 120s. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when Mauchly's test showed a violation of sphericity.

Effect	Normalized Maximum Lyapunov Exponent			
	df 1,2	F	p	η_p^2
Driving Condition (DC)	**2.185, 349.573**	**5.301**	**0.004**	**0.032**
Sources (S)	**1,160**	**42.269**	**<0.001**	**0.209**
Gender (Gn)	1,160	0.270	0.604	0.002
Group (G)	1,160	1.695	0.195	0.010
Participants (P)	7,160	1.956	0.064	0.079
Based Interval (Based)	1,160	0.258	0.612	0.002
DC x S	**2.185, 349.573**	**5.319**	**0.004**	**0.032**
DC x P	**8.739, 349.573**	**2.117**	**0.029**	**0.050**
DC x Based	2.185, 349.573	1.953	0.139	0.012
DC x S x P	8.739, 349.573	0.937	0.491	0.023
DC x S x Based	2.185, 349.573	1.120	0.331	0.007
DC x P x Based	8.739, 349.573	1.028	0.416	0.025
DC x S x P x Based	8.739, 349.573	1.337	0.218	0.032

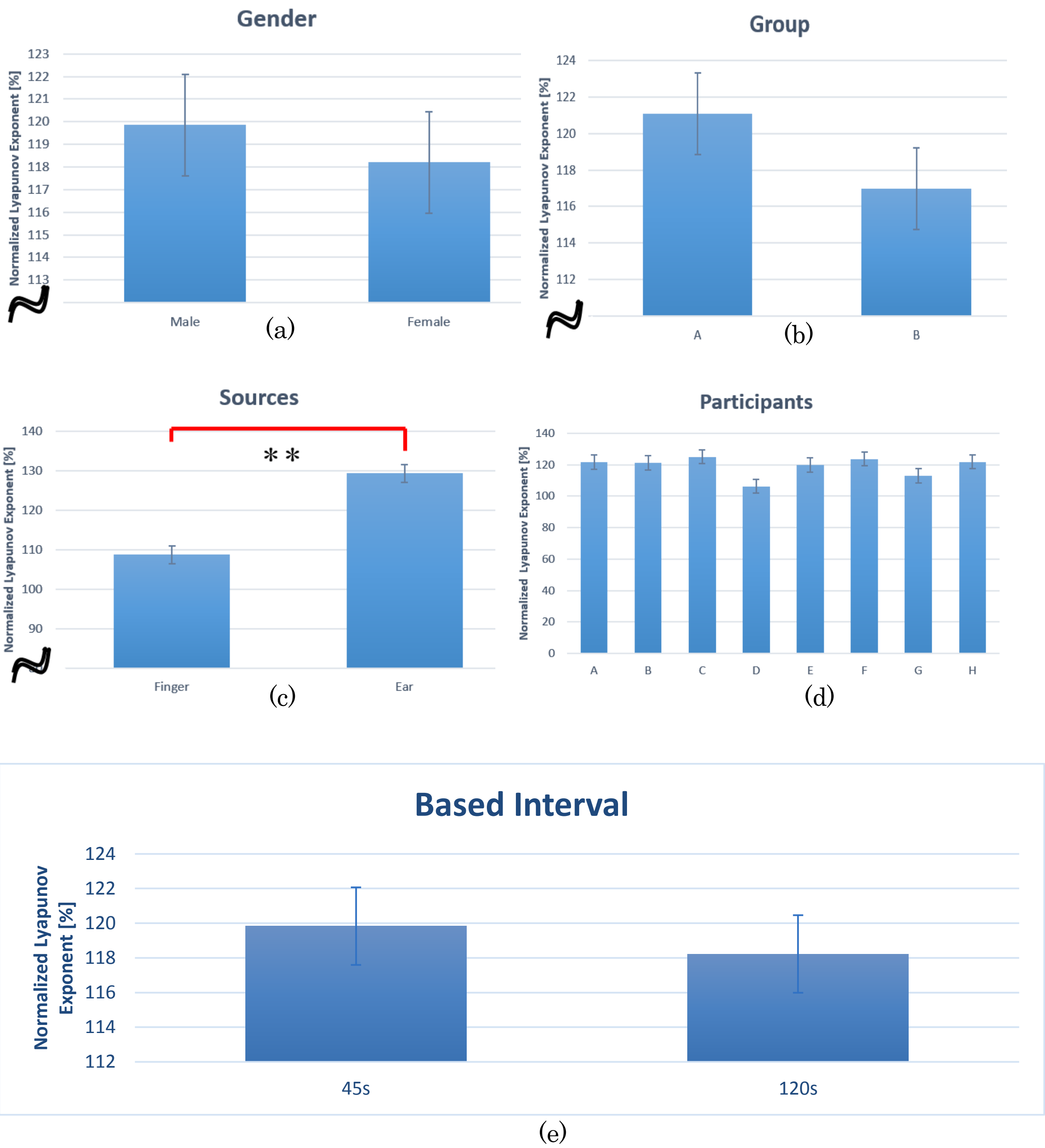


Figure 3.13 Univariate test results for Normalized Lyapunov exponent

Table 3.6 and Figure 3.13 showed the results of the univariate test for the Normalized Lyapunov exponent. From the table, we evaluate the main effect as follows;

<u>Gender</u>

According to the analysis, the main effect of gender was not significant, $F(1,160)=0.270$, $p=0.604$, $\eta_p^2=0.002$.

Experimental Order

As mentioned in the method section, we divided the participants into two groups with the different between the orders of the experiment. According to the analysis, the main effect of group was not significant, $F(1,160)=1.695$, $p=0.195$, $\eta_p^2=0.0.10$.

Types of Base interval

According to the analysis, the main effect of types of the base interval was not significant, $F(1,160)=0.258$ $p=0.612$, $\eta_p^2=0.002$.

Individual differences between participants

Based on the results of analysis, the individual differences between participants were also not significant, $F(7,160)=1.956$, $p=0.064$, $\eta_p^2=0.079$.

The interaction between Driving condition and Sources, $F(2.2,349.6)=5.319$, $p=0.004$, $\eta_p^2=0.032$, between Driving conditions and Participants, $F(8.7, 349.6)=2.117$, $p=0.029$, $\eta_p^2=0.05$, were found significant.

Based on the results and comparisons between the average value of maximum Lyapunov exponent and Normalized Lyapunov exponent, we come to a conclusion that normalization of data was able to remove the individual differences between participants such as between gender, between two groups of difference experimental order, and between individual differences of data for 8 participants. It is worth to do the normalization as this will further increase the accuracy of estimation of mental workload level in our study.

To further discuss between the selections of based interval, as we found no significant difference between base intervals in the previous analysis, we tried to separate the analysis of the sources of data. This process was done due to significant differences found between the sources, $F(1,160)=42.269$, $p<0.001$, $\eta_p^2=0.209$. With that consideration, analysis of selection between the two based intervals was performed only on data from Ear BVP sensor.

3.4.1.2.1 Normalized Maximum Lyapunov exponent 45s (Base interval: 45s~90s)

The result for Normalized Maximum Lyapunov exponent for 8 participants was illustrated in Figure 3.14 and Table 3.7. As mentioned above, the data were taken only from Ear BVP sensor. A repeated-measures ANOVA was conducted with driving conditions as a within-subjects factor. The results show that there was a significant main effect of driving conditions on the value of Maximum Lyapunov exponent ($F(5, 43) = 5.802$, $p<0.000$, $\eta_p^2=0.403$). Further analysis of pairwise comparisons using Bonferroni Method revealed that significant differences were found between, NMATBD1 and MAT1HC ($p=0.005$), between NMATBD1 and MAT2HC ($p=0.003$). The differences between NMATBD2 and MAT1HC ($p=0.03$) and NMATBD2 and MAT2HC ($p=0.001$) were also significant.

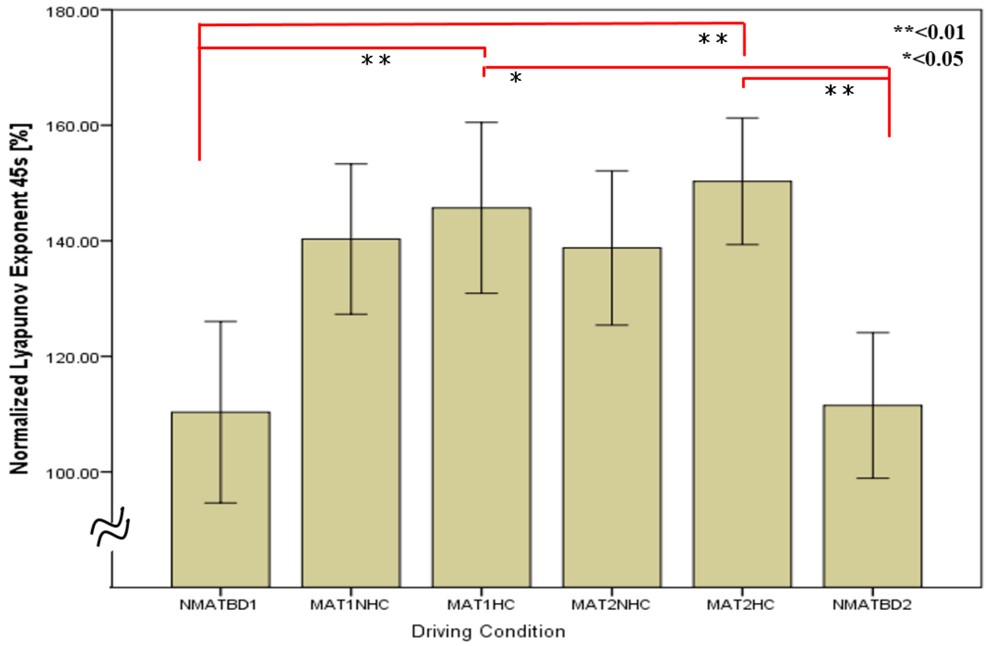


Figure 3.14 Average value of Normalized Lyapunov Exponent 45s

When looking at the value, the day when participants conducted MAT2 under HC of traffic condition showed the highest value of the Normalized Maximum Lyapunov exponent with 150.30%. This data indicated that the participants were in a high mental workload when performing this type of task under a hazardous condition.

Overall, the values of Normalized Lyapunov exponent 45s for the days with secondary tasks were higher than the values of the days without a secondary task. This finding suggested that it is possible to distinguish the value of Normalized Lyapunov exponent considering the days with and without the secondary task.

For the level of difficulties of tasks, the differences between the driving conditions with the secondary task and traffic conditions were not significant.

Table 3.7 Pairwise comparisons for Normalized Lyapunov Exponent 45s. Significant effects ($p<0.05$) are shown in bold.

	NMATBD1	MAT1NHC	MAT1HC	MAT2NHC	MAT2HC	NMATBD2
NMATBD1	X	0.061	**0.005**	0.225	**0.003**	1.000
MAT1NHC	0.061	X	1.000	1.000	1.000	0.123
MAT1HC	**0.005**	1.000	X	1.000	1.000	**0.03**
MAT2NHC	0.225	1.000	1.000	X	1.000	0.067
MAT2HC	**0.003**	1.000	1.000	1.000	X	**0.001**
NMATBD2	1.000	0.123	**0.03**	0.067	**0.001**	X

3.4.1.2.2 Normalized Maximum Lyapunov exponent (Base interval: 60s~120s)

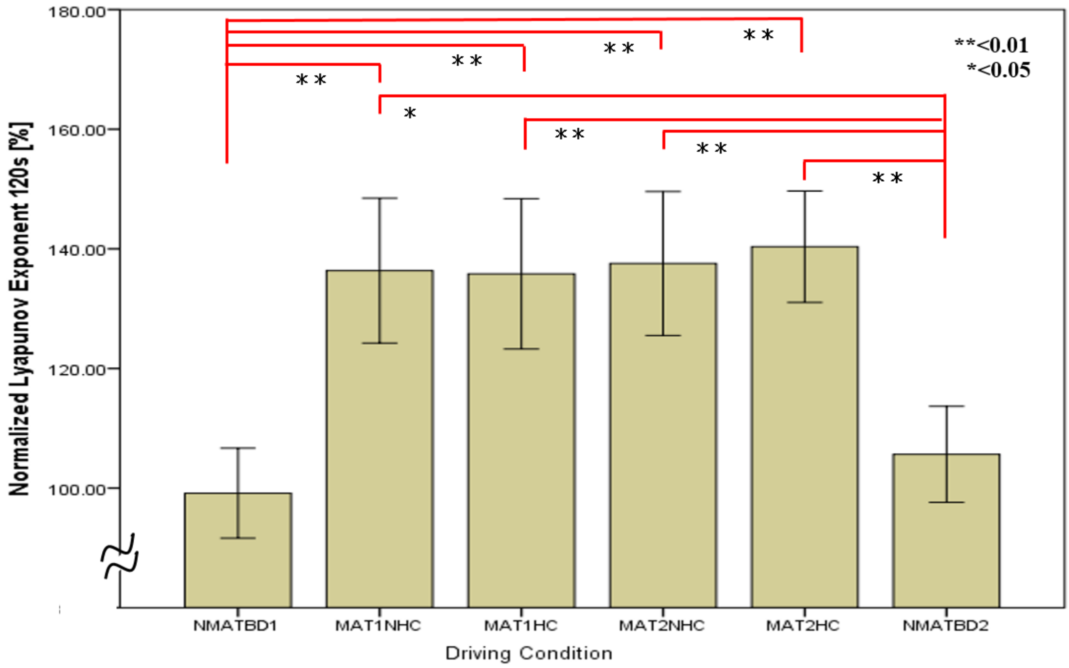


Figure 3.15 Average value of Normalized Maximum Lyapunov exponent 120s

Figure 3.15 shows the result of Normalized Maximum Lyapunov exponent with a base interval of 60s~120s.Table 3.8 showed the result of repeated measures ANOVA among the days of driving conditions. A significant main effect was found between the Driving Conditions ($F (5, 43) = 13.272$, $p=0.000$, $\eta_p^2 =0.607$).

Table 3.8 Pairwise comparisons for Normalized Lyapunov Exponent 120s. Significant effects (p<0.05) are shown in bold.

	NMATBD1	MAT1 NHC	MAT1 HC	MAT2 NHC	MAT2 HC	NMATBD2
NMATBD1	X	**0.000**	**0.000**	**0.000**	**0.000**	1.000
MAT1 NHC	**0.000**	X	1.000	1.000	1.000	**0.011**
MAT1 HC	**0.000**	1.000	X	1.000	1.000	**0.009**
MAT2 NHC	**0.000**	1.000	1.000	X	1.000	**0.000**
MAT2 HC	**0.000**	1.000	1.000	1.000	X	**0.000**
NMATBD2	1.000	**0.011**	**0.009**	**0.000**	**0.000**	X

As shown in the table, a significant difference was found between NMATBD1 and driving conditions where they accomplished the tasks. The differences between the NMATBD2 and other driving conditions were also significant. The highest value was reported from the MAT2HC driving condition, where they have to accomplish the difficult level of MA Task under Hazardous traffic condition. As the value in this condition task was the highest among the other days, it is possible to estimate the highest mental workload of participants. Same with the previous observation in the average maximum Lyapunov exponent and the Normalized maximum Lyapunov exponent 45s with base interval 45s~90s, the participants were in the highest mental workload when performing the MAT2 task under HC traffic condition.

The value of Normalized Maximum Lyapunov exponent for the days with a secondary task was also recorded as higher compared to the days without a secondary task. For the levels of difficulties of a task, there was no significant difference between the driving conditions with a task under either hazardous condition or Non-Hazardous traffic condition.

3.4.2 Performance measures

3.4.2.1 Primary Task

Standard Deviation of Lateral Position

As the maintaining safety of vehicle was regarded as the primary task, maintaining the lateral position for not deviating from the center of the cruising lane is a good indicator for the measurement. Lateral position is the vehicle's position in relation to the center of the lane in which they are driving. Figure 3.16 shows a part of the driving track. In this experiment, the lane width was 3.3m. The standard deviation of Lateral position (SDLP) of the vehicle was calculated within the MAT interval as an indicator of the degree of swerving or deviating from the center line.

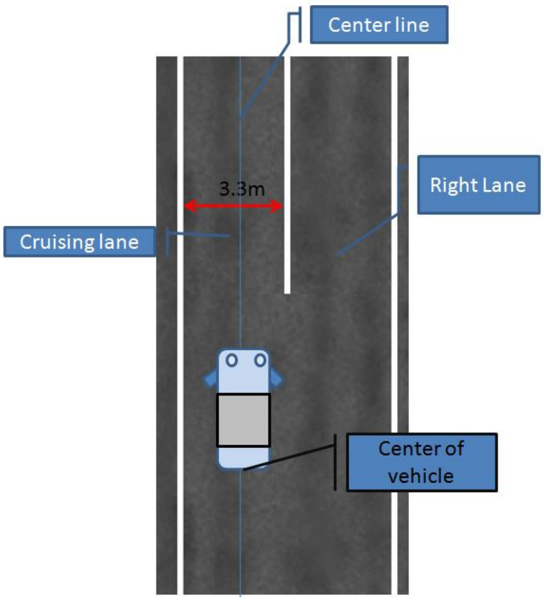


Figure 3.16 Driving track and center line of cruising lane

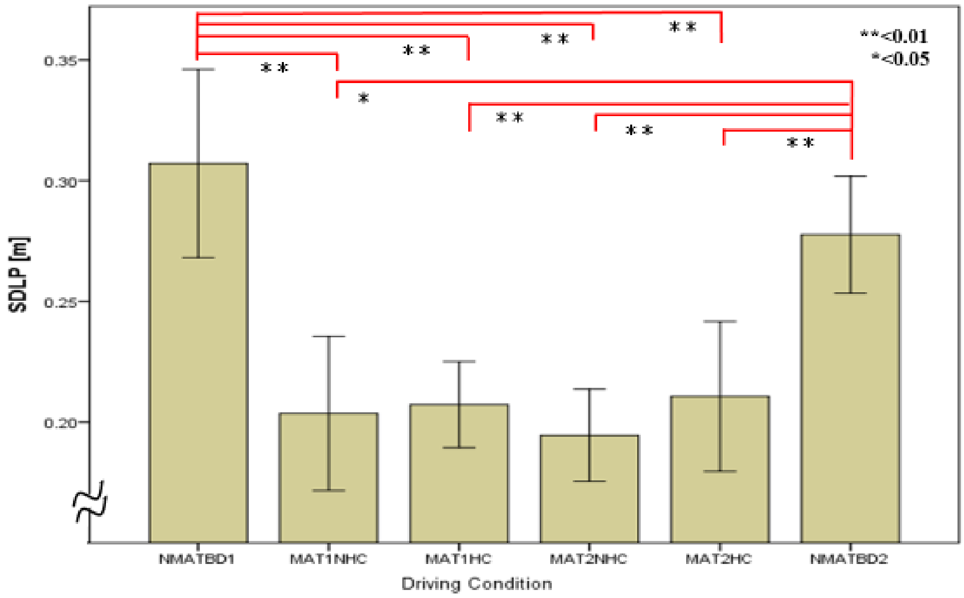


Figure 3.17 Average of Standard Deviation of Lateral Position

In Figure 3.17, the average of the standard deviation of lateral position (SDLP) for 8 participants is shown. Compared to the baseline driving a decrease in SDLP was found in conditions in which participants were performing a secondary task, (MAT1NHC, MAT1HC, MAT2NHC, and MAT2HC). The highest value of SDLP was found on the first day of an experiment which was the baseline driving. It could be explained that as driving on NMATBD1 was the first day of the experiment, participants needed time to be accustomed to the driving simulator, and this resulted the SDLP was the highest.

A repeated-measures ANOVA was performed, and the results showed that the main effect of driving conditions was significant on SDLP ($F (5, 43) = 8.334$, $p<0.001$, $\eta_p^2 = 0.492$). The results of pairwise comparisons using Bonferroni method in Table 3.9 showed that except for NMATBD1 and NMATBD2, there was no significant difference between the other task. As there were no significant differences between MAT1NHC, MAT1HC, MAT2NHC and MAT2HC, the decrement of SDLP was maybe because of participants were accomplishing the secondary task under either hazardous or non-hazardous traffic conditions.

Table 3.9 Post Hoc Test using Bonferroni Method for average SDLP. Significant effects (p<0.05) are shown in bold.

	NMATBD1	MAT1NHC	MAT1HC	MAT2NHC	MAT2HC	NMATBD2
NMATBD1	X	**0.001**	**0.001**	**0.000**	**0.006**	1.000
MAT1NHC	**0.001**	X	1.000	1.000	1.000	**0.018**
MAT1HC	**0.001**	1.000	X	1.000	1.000	**0.000**
MAT2NHC	**0.000**	1.000	1.000	X	1.000	**0.000**
MAT2HC	**0.006**	1.000	1.000	1.000	X	**0.015**
NMATBD2	1.000	**0.018**	**0.000**	**0.000**	**0.015**	X

The decreasing value of SDLP while accomplishing the task was not something that we expected. Our expectations were the more difficult the task is, the more SDLP value could be observed as performance deteriorate.

3.4.2.2 Secondary Task

For secondary task performance, we measured the percentage of the correct answer on mathematical arithmetic tasks that participants performed. The percentage of the correct answer is a percentage of the times when participants answer was correct. The lower percentage of correct answer indicates a higher mental workload.

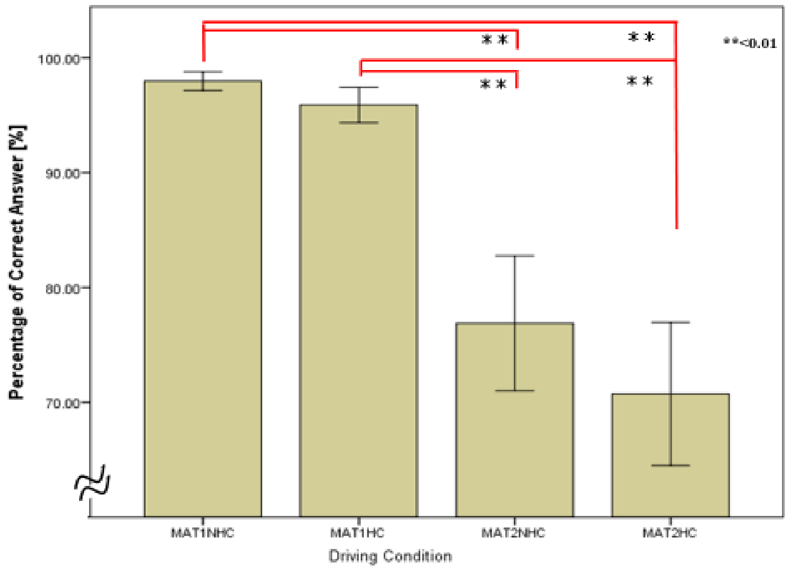


Figure 3.18 Percentage of correct answer on MAT

Figure 3.18 showed the percentage of correct answers for MA task. The percentage of correct answers for MA task decreased as the participant performed a harder task. A repeated-measures ANOVA was conducted, and the results showed that the main effect of driving conditions was found on the percentage of the correct answer ($F(3, 53) = 32.964$, p=0.000, η_p^2 =0.687). A posthoc test using Bonferroni method was performed and the results as shown in Table 3.10.

Table 3.10: Pairwise comparisons for a percentage of the correct answer. Significant effects (p<0.05) are shown in bold.

	MAT1NHC	MAT1HC	MAT2NHC	MAT2HC
MAT1NHC	x	0.059	**0.000**	**0.000**
MAT1HC	0.59	x	**0.000**	**0.000**
MAT2NHC	**0.000**	**0.000**	x	0.242
MAT2HC	**0.000**	**0.000**	0.242	x

The lowest percentage of the correct answer was recorded at an MAT2 task where the drivers were required to solve two digit numbers of the mathematical arithmetic task under a hazardous traffic condition.

The result of pairwise comparisons between driving conditions indicated that the significant differences were found between MAT1 and MAT2 (Table 3.). The difference between traffic conditions was not significant.

It can be concluded the more difficult the secondary task was, the performance of the task decreased.

3.4.3 Subjective measurements

After completing each trial, all participants were requested to estimate the workload they experienced during the trial. Figure 3.19 showed an average rating for mental workload from NASA-TLX for all eight participants.

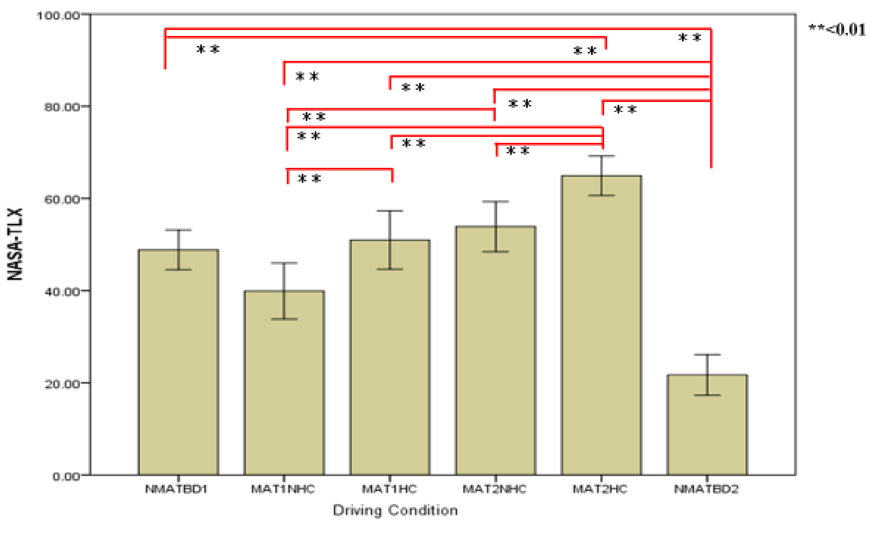


Figure 3.19 Results of NASA-TLX

A repeated-measures ANOVA was performed, and the results showed that the main effect of driving conditions on the NASA-TLX was significant (F (5, 43) = 45.921, p=0.000, η_p^2 =0.842).

Table 3.51 Post Hoc Test results. Significant effects (p<0.05) are shown in bold.

	NMATBD1	MAT1NHC	MAT1HC	MAT2NHC	MAT2HC	NMATBD2
NMATBD1	X	0.071	1.000	0.878	**0.000**	**0.000**
MAT1NHC	0.071	X	**0.000**	**0.000**	**0.000**	**0.000**
MAT1HC	1.000	**0.000**	X	**0.000**	**0.000**	**0.000**
MAT2NHC	0.878	**0.000**	1.000	X	**0.000**	**0.000**
MAT2HC	**0.000**	**0.000**	**0.000**	**0.000**	X	**0.000**
NMATBD2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	X

Table 3.51 showed the result of Post-Hoc test results for NASA-TLX using Bonferroni Method. The average value of mental workload was the highest in MAT2HC with a mean value of 66.54 %. It has to mentioned here that there was a significant different between the first (NMATBD1) and the last day (NMATBD2) of experiments. The participants may have evaluated the first day of the experiment as in a high mental workload since they need to be accustomed to the driving simulator. For the last day of the experiment, after experiencing all of the tasks, and accustomed with the driving simulator, they evaluated the mental workload as the lowest.

3.5 Discussion

3.5.1 Normalization data

In this experiment, one of the objectives was to distinguish between the level of difficulties of tasks by combining between internal and external factors in increasing mental workload. To evaluate the mental workload and finding a standard value of mental workload, we attempted to normalize the physiological measurements data (in this experiment, maximum Lyapunov exponent). When considering the normalization which we defined as comparing the MAT interval with a base interval as a reference, we tried to give an insight on selecting between the base interval of 45s and 120s. The results of normalization with repeated-measures ANOVA shows that the normalization was able to reduce the differences such as between Gender, between orders of experiments with different groups and between participants themselves.

Based on the results of normalizations, we did not find any significant difference between selecting based interval either 45s or 120s. Based on further observation of the results, we also found that the sources of data were a significant difference between Ear or Finger data. This result means that the data between ear and finger was totally different when normalizing the data. To give an insight regarding selecting these two base intervals, we further analyzed the data with selecting only data of ear BVP sensor. We further analyzed the normalization data with a repeated-measures ANOVA and driving conditions as within-subjects factors. We found that compared to Normalization 45s, Normalization 120s was able to give significant differences driving conditions. Such as between NMATBD1 and MAT1NHC, between NMATBD1 and MAT2NHC, also between NMATBD2 and MAT1NHC and NMATBD2 and MAT2NHC. Although the differences only found for these driving conditions we were able to distinguish between selecting the two base intervals. Overall, it gives an insight that normalized Lyapunov exponent 120s were able to make a distinction between the driving conditions where participants were performing the tasks and the baseline driving (NMATBD1 and NMATBD2). In other words, through normalization of 120s, it might be possible to make an estimation of driver's mental workload between a normal state, when the participants were not performing any task and performing either MAT1HC, MAT1HC, MAT2NHC, or MAT2HC. We can confirm that there were two levels in the tasks.

3.5.2 Cross-Evaluation between measurement groups

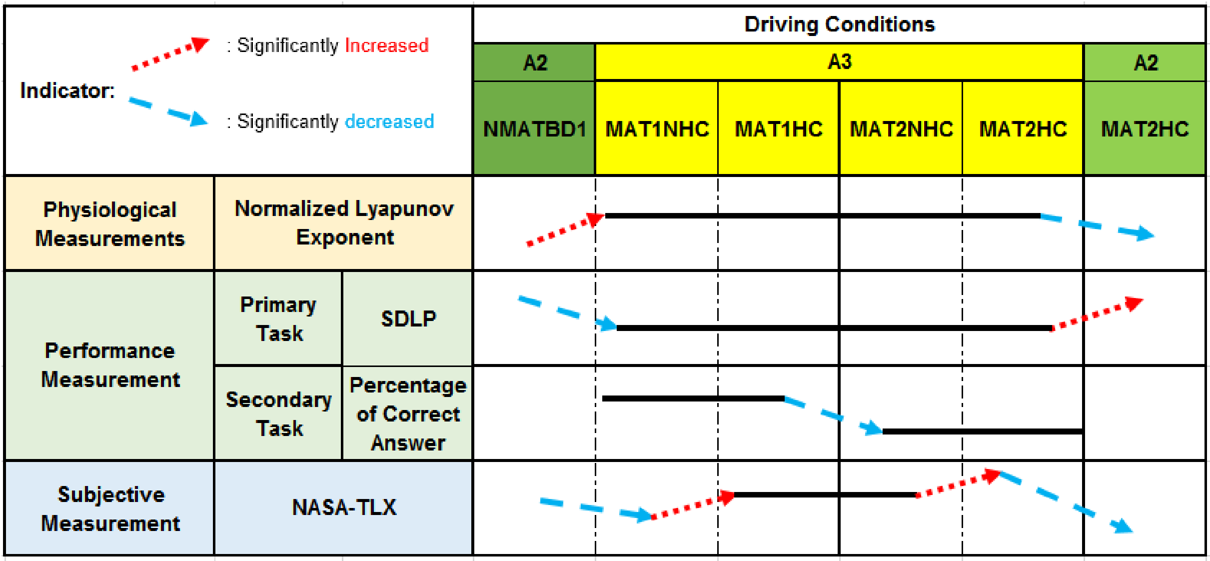


Figure 3.20 Overall results for evaluation of mental workload.

Figure 3.20 shows an overall result for evaluation of mental workload. The red dotted line shows a significantly increased on the data while the blue dotted line shows a significantly decreased. By looking at this figure which mapping together all the measurements, we can obtain a clearer picture on the results. This kind of information can lead to a better interpretation on the mental workload estimation.

As shown in the figure, the primary task performance of SDLP was decreased when participants were accomplishing the tasks. This result was contradictory to our expectation, as we predict that more difficult the task is performance will degrade. This expectation is based on single resource theory by Kahneman [21] and also the conceptual theory by de Waard [20] regarding the relationship between task demand and performance. One of the explanations for this circumstance was they increased effort to cope with the increasing demand of mental workload. This notation can be seen by looking at the deterioration of secondary task performance. The increasing effort on the task and improvement of performance also can be regarded as "tunnel vision" or a narrowing of attention. This phenomenon can be regarded as improvement in lateral position control [2], [49], [51], [72–76].

As we performed normalization of physiological measurement data, we can map together this measurement with a measurement from a different group which is subjective measurements. This step also one of the processes of cross-evaluation to get a higher accuracy on mental workload estimation.

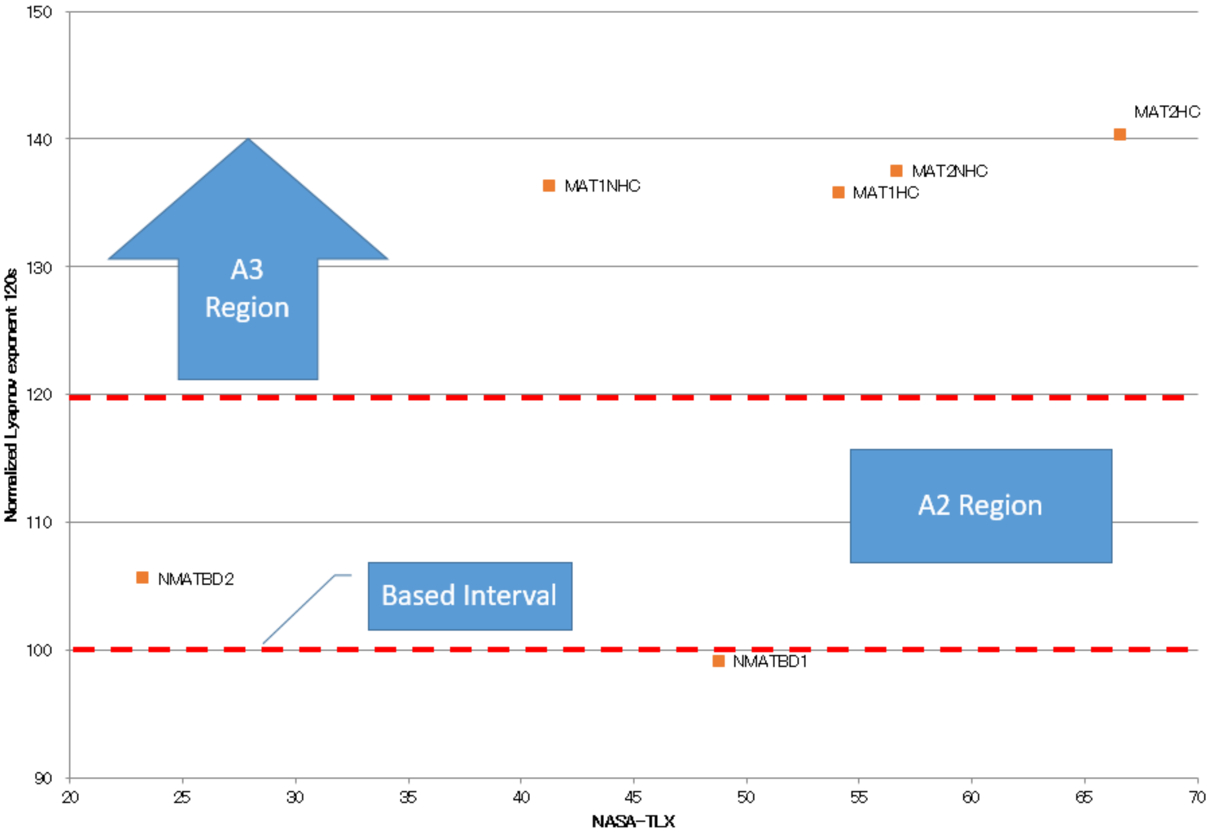


Figure 3.21 Mapping graph between NASA-TLX and Normalization Lyapunov exponent 120s

Figure 3.21 showed a mapping graph between the subjective rating of mental workload and Normalized Lyapunov exponent 120s. As shown in the figure, horizontal line of 100% indicates the based interval used in the normalization. As we can see, NMATBD2 and NMATBD1 were around the based interval indicate that data were not much difference with the baseline period. By using this information, we tried to map the finding with de Waard`s model as in Figure 2.2. For this experiment, we were able to distinguish two regions, Region A2 where the mental workload was low and Region A3 where the participants were trying their hard in maintaining the performance while the mental workload was increased. We think it is appropriate to distinguish between these two regions by drawing another horizontal line on 120%. This value is based on a consistent finding by the study of Suzuki and Okada [71] who draw the threshold of 120% when the drivers were facing the hazardous situation.

At the beginning of this experiment, we also hypothesized that the most difficult task by combining the difficulty of the external and internal factors (in this experiment: MAT2HC) would give the highest value of mental workload. It appeared to be true for the subjective measurement and secondary task performance. However, for the Physiological measurement, we need another explanation for the result of there was no significant differences between the driving conditions except for between the baseline. We further designed Experiment#2.

3.5.3 Factors of increasing mental workload

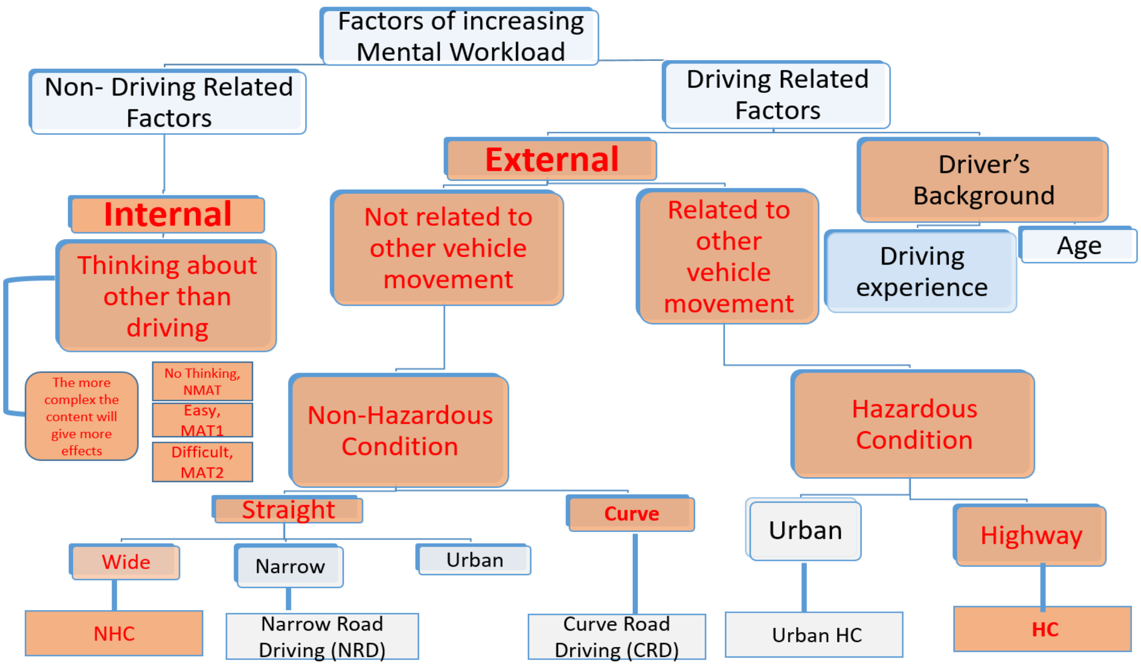


Figure 3.22 Factors of increasing mental workload in Experiment#1

Figure 3.22 shows a diagram on the factors of increasing mental workload in Experiment#1 that we think would affect the results. The red font and column show the factors that we think contribute to increasing mental workload in Experiment#1. As we can see, various factors give effects to increasing mental workload in Experiment#1. We will try to focus on answering the question regarding the hypothetical relationship between performance and driving performance by considering internal and external factors by designing Experiment#2.

3.6 Conclusions

These are the findings we obtained from this experiment;

1. The normalization of physiological measurement data was able to remove individual differences of data.
2. Normalized Lyapunov exponent 120s which used 120 seconds before starting the MAT as a based interval was able to distinguish most driving conditions than normalized Lyapunov exponent 45s.
3. By combining between Normalized Lyapunov exponent 120s and NASA-TLX, we were able to differentiate between normal driving (Region A2) and during the participants accomplishing the tasks (Region A3), either external or internal factors of mental workload.
4. By using a cross-evaluation method between all measurement groups, we can obtain a better interpretation on the mental workload estimation.
5. While accomplishing the tasks, the primary task performance showed a decrease in the SDLP.

Chapter 4 Internal Factors of Mental Workload

4.1 General overview and aims

After achieving the key findings in the method of increasing the accuracy of estimation in Experiment#1, we further investigate the level of difficulties in internal and external factors of mental workload. As mentioned in the introduction, some of the important questions we want to answer are, “Is there exist interaction between internal and external factors?”, “If the interaction exists, which one gives more effect to driver’s mental workload?”, And “How can we improve our understanding of the relationship between mental workload and driving performance?”. To answer these questions, we designed Experiment#2 based on the results we obtained from Experiment#1. In this experiment, we also evaluate the usage of a Blood Volume Pulse sensor as a potential measurement to measure mental workload.

The specific aims of this experiment as follows:

- To validate the hypothetical relationship between mental workload and performance
- To acquire a better understanding of mental workload theory under consideration of internal and external factors.
- To evaluate the effects of **the internal factors** of mental workload to driver’s mental workload.
- To evaluate the usage of BVP sensor as a potential mental workload measurement.

4.2 Hypotheses

Our hypotheses for this experiment as follows:

1. Blood Volume Pulse (BVP) sensor can be used to evaluate the driver’s mental workload.
2. The internal factors such as thinking other than driving give more effects to the mental workload measurements.
3. The increasing value of mental workload measurements come from **internal and external** factors.
4. The decreasing value of SDLP does not mean the improvement in performance.
5. The more complex the content of thinking will give more effects to the mental workload measurements.

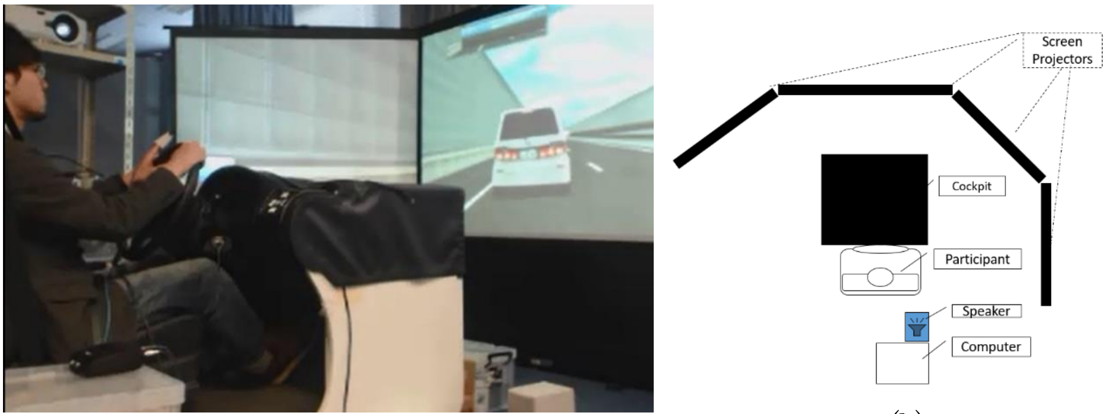
4.3 Method

4.3.1 Participants

Twenty-five male drivers participated in the experiment. Their mean age was 21.1 years old, and the standard deviation was 1.3. Every participant held a valid driver's license. The average driving experience was 2.3 years. They were paid off 820yen per hour, and the total length of the experiment was 3 hours.

4.3.2 Apparatus

This study used a fixed-base driving simulator developed by Mitsubishi Precision, Inc (Figure 4.1 (a)). It consists of 4 screens (H120 x W160mm) been projected by projectors. The location of screen projectors as shown in Figure 4.1 (b). Blood Volume Pulse (BVP) data were recorded with the Nexus-10 device. The software on Mind media was used to analyze data of HR and HRV. BVP device attaches to a finger of participants.



(a) (b)

Figure 4.1 The driving simulator used in the study

4.3.3 Tasks

4.3.3.1 Primary Task (External Factors)

The primary task was to drive safely in the left lane. The driving course was designed to be a straight pathway with approximately 10km distance with 3.5m wide lane. Every participant experienced two levels of difficulties of traffic conditions, a) None Hazardous Conditions (NHC) and b) Hazardous Condition (HC). Under both conditions, participants were asked to follow a lead vehicle (LV).

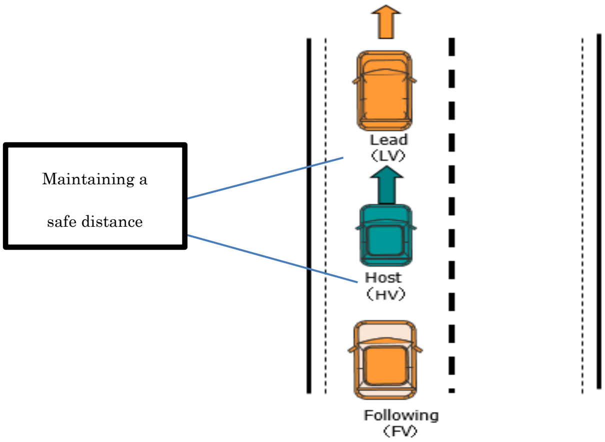


Figure 4.2 Non-Hazardous Condition (NHC)

Figure 4.2 describes how the driving task was given to each participant in NHC. The Following Vehicle (FV) was located behind the HV to help participants maintain a following distance. In this traffic condition, FV drove with 90km/h (constant speed) throughout the 6-minute trial. Some vehicles also existed in the right lane.

Under HC condition, participants were also asked to maintain the distance between LV and FV. At this time, both LV and FV cruised between the speed of 55km/h and 100km/h. Intentionally LV and FV performed an abrupt braking task of 0.35g and also a quick acceleration. Time to make a sudden braking and rapid acceleration were determined at random (approximately twice the speed of the changes for each 400-meter run). Overall the LV made a braking task 15 times including two dummy braking tasks (LV activated the braking lamp yet it accelerate) throughout 6-minutes of a trial. The dummy braking tasks were designed so that participants aware of the situations (for not just depending on the braking lamp on Lead Vehicle to maneuver the brake pedal). The participants were requested to ensure a safe following distance and to be careful to an abrupt variation of speed on both FV and LV (Figure 4). If they meet with a crash during a trial, participants would have to start a new trial all over again. This way, participants will try their best for not being involved in a crash.

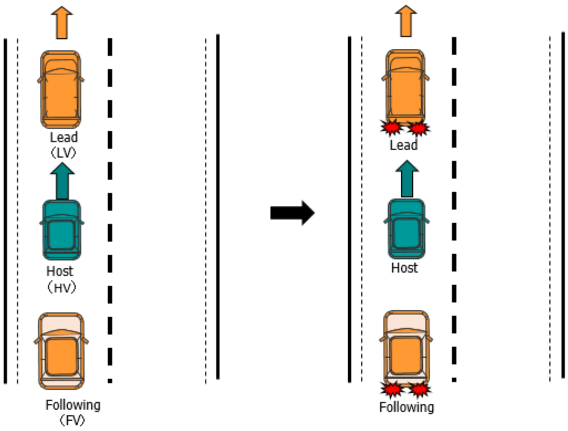


Figure 4.3 Hazardous condition (HC)

4.3.3.2 Secondary Task (Internal Factors)

For the internal factor of mental workload, a secondary task has been used. Participants were requested to carry out a two-minute Mathematical Arithmetic Task (MAT) in a 6-minute trial, two minutes after the start and two minutes before the trial completed (see Figure 4.4).

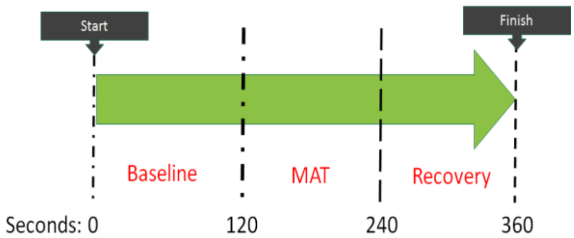


Figure 4.4. MAT period in a trial

MAT requires the participants to memorize the numbers they heard three seconds before and at the same time try to solve the arithmetic. This task is a kind of so-called PASAT (Paced Auditory Serial Addition Test) [69]. The mathematical arithmetic task was divided into two levels, namely the level of easy (MAT1) and the level of difficulty (MAT2) tasks. In MAT1, participants were given single-digit numbers (from 1 to 9) in every three seconds through a speaker connected to a computer (Figure 4.1b). Participants had to give the answer of the summation of the last two numbers verbally as in Figure 6. While in MAT2, participants were given two-digit numbers between 11 and 49. As in MAT1, the participants had to answer a total of the last two numbers orally.

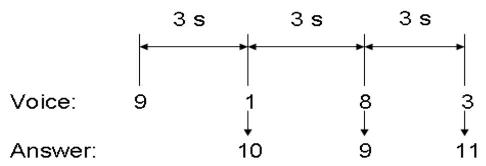


Figure 4.5 MAT1 (Easy Level)

While performing a mathematical task, it was assumed that demand of the mathematical task had remained unchanged during the 2-minute task. In each trial, the number of the correct answers was calculated, and the results had been informed to the participants after the end of the trial. We thought that by knowing the results, they would try their best to attain a good result during the secondary task.

Non-Driving Task (NDT)

In this situation, participants had to seat on driving seat and wait until two minutes to perform either MAT1 or MAT2. NDT had been set before start the experiment with driving conditions to test on how the MAT alone gave an effect to the measurements.

4.3.4 Experimental Design

The experiment was designed using two manipulations in a within-subject factor factorial design: Secondary Task (NMAT, MAT1 or MAT2) and Traffic Condition (NHC or HC). To cancel the possibility of existing an order effect, the experimental procedure was arranged randomly from NMATBD1 to NMATBD2 as in Table 1. BD was 'Baseline Driving' where they had been asked to follow LV with a constant speed (85km/h) and without FV on behind. In this situation, there was no secondary task imposed to the participants.

Table 4.1. Experimental procedure

Randomly arranged

No. of Trials	Secondary Task	Traffic Condition	Driving Condition
1	NMAT	NHC	NMATNHC
2	NMAT	HC	NMATHC
3	MAT1	NHC	MAT1NHC
4	MAT1	HC	MAT1HC
5	MAT2	NHC	MAT2NHC
6	MAT2	HC	MAT2HC

Flow of Experiment
Practise-1
Practise-2
Rest
MAT1NDT
MAT2NDT
NMATBD1
Trial 1
Rest
Trial 2
Trial 3
Trial 4
Rest
Trial 5
Trial 6
NMATBD2

We expected that both external and external factors of mental workload gave an effect to the measurement groups of mental workload. It was also predicted that driver workload was higher during the combination of the difficulty level of an internal factor of workload under the difficulty level of the external factor. We also expected to find Region A2 where task-related effort while the driver tries their hardest to maintain the performance.

4.3.5 Procedure

Upon arriving at the venue of the experiment, participants were explained regarding the experiment and signed a consent letter. They were given time to get familiarized with the simulator and also the MAT. Before starting the experiment, BVP device was placed on either right or left hand (which one not dominant) finger of participants. As shown in Table 4.1, every participant performed six types of combination between secondary task and the traffic condition. As a subjective measurement, participants were requested to answer NASA-TLX [27] and RSME [64] immediately after finishing every trial.

4.3.6 Dependent variables

Table 4.2 List of Dependent Variables

Measurement Groups	Dependent Variables
Physiological	Heart Rate (HR) and Heart Rate Variability (HRV) data from Finger BVP sensor
	Heart Rate (HR) and Heart Rate Variability (HRV) data from ECG
Primary Task Performance	Standard Deviation of Lateral Position
	Steering Entropy
Secondary task Performance	Percentage of Correct answer
Subjective	NASA-TLX
	RSME

Table 4.2 shows the list of dependent variables tested in this experiment. HR is the speed of the heartbeat measured by the number of the beat of the heart per minute (Beat-Per-Minute, BPM). BioTrace software was used to analyze the HRV data. HRV data can be analyzed by using time-domain and frequency domain. For this research, we used frequency domain. First, the continuous BVP data were transformed to Inter-Beat-Interval (IBI) time series. Then by using Fast Fourier Transform (FFT) Spectral Analysis, IBI converted to the frequency domain of HRV Spectrum. From the HRV Spectrum, there are two bands of interest: High Frequency (HF) band (0.15Hz - 0.7Hz) and Low Frequency (LF) band (0.04 Hz - 0.15Hz) [77-78]. We used the ratio of LF/HF as it can measure of Autonomic Nervous System balance.

4.3.7 Data analysis

To increase the accuracy of estimation the levels of mental workload as been discussed in chapter 3, we normalized physiological and performance data. This method is useful to find a standard value that indicates the same level of mental workload. It was also to ensure that data were not affected by "state-related effort" especially fatigue, e.g. considering the data maybe has already increased from the beginning of one trial and not only during performing MAT.

Normalization equation as shown in equation (1). In a 6-minute trial, minute 2 to minute 4 from the beginning was regarded as MAT Period data.

We set two minutes at the start of a trial as a 'Base Interval' and compared the data to MAT period from minute 2 to minute 4 of a trial. Results from Experiment#1 were used to determine the base interval which is 120s (the same length with MAT period data). The interval between minute-0 to minute-2 was selected as base interval data.

$$Normalization = \left(\frac{MAT\ Period\ Data}{Base\ Interval\ Data} \times 100\right)\% \tag{1}$$

General Linear Model for repeated-Measures analysis in SPSS (version 21.0) was used to analyze the data. Two independent variables: type of traffic conditions (NHC and HC) and type of secondary Task (NMAT, MAT1, and MAT2) were analyzed by using a 2 X 3 repeated-measures design.

4.4 Results

4.4.1 Ordering effects of experimental

To test whether the ordering effects existed on the measurements, we first analyzed data of NMATBD1 and NMATBD2.

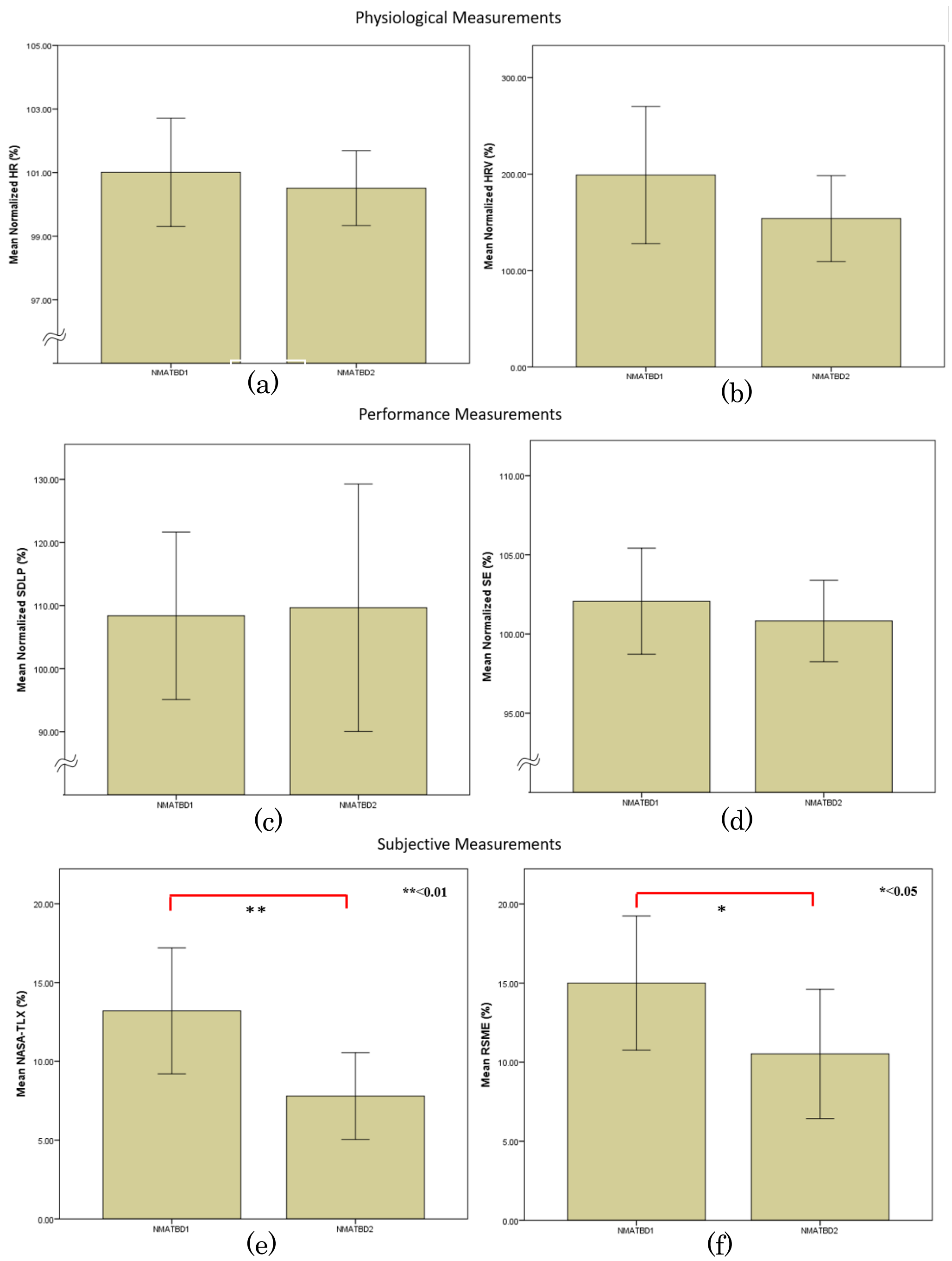


Figure 4.6(a)~(f): Results of mean value of NMATBD1 and NMATBD2 for each measurement

Figure 4.6 (a)~(f) show the results of the average value of NMATBD1 and NMATBD2 for each measurement. A paired-samples t-test was conducted to compare between NMATBD1 and NMATBD2. There was no significant difference in the data from HR for NMATBD1 (M = 101.01, SD=4. 13) and NMATBD2 (M=100. 51, SD=2. 85); t (24) =0. 544, p=0. 592). There was also no significant difference in the data from HRV for NMATBD1 (M=198.96.08, SD=172.08) and NMATBD2 (M=153.87, SD=107.87); t(24)=1.13, p=0.269. These results suggest that the tasks they performed during the experiment did not give an order effect to physiological measurements for the first and the last order of driving tasks.

For performance measurements, there was no significant difference found between NMATBD1 (M=108. 37, SD=32. 166), NMATBD2 (M=109. 65, SD=47. 49) at t(24) =. 122, p=. 904 for the SDLP. The differences also not significant for Steering Entropy between NMATBD1 (M=100.02, SD=0.08) and NMATBD2 (M=100.00, SD=0.06) t(24)=0.628,p=0.536.

While for subjective measurements there were significant differences between NMATBD1 and NMATBD2. Results from NASA-TLX shows a significant difference between NMATBD1 (M=13.20, SD=9.69), NMATBD2 (M=7.80, SD=6.67), t(24)=3.441, p=0.002. And the results from RSME also revealed the same pattern for NMATBD1 and NMATBD2, NMATBD1 (M=15.00, SD=10.27), NMATBD2 (M=10.52, SD=9.91), t(24)=2.531, p=0.018.

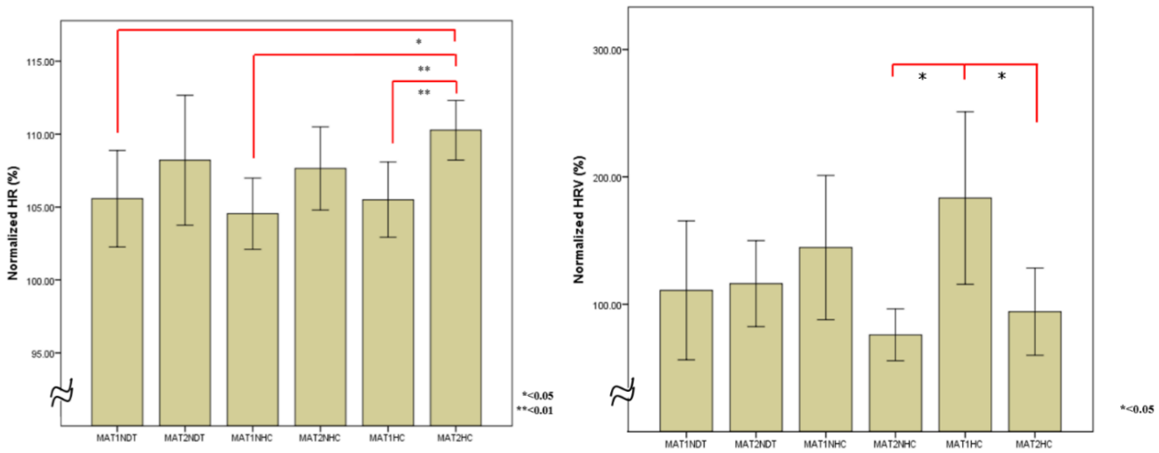
These results of NMATBD1 and NMATBD2 suggest that there were no ordering effects for performance and physiological measurements, however, participants rate differently between the first and the last run of an experiment for subjective measurement. They might feel there were differences between the first and the last order of the driving task because they felt that they were not accustomed to the tasks for the first time. However, the results for performance and physiological measurement displayed differently as they were no significance difference between NMATBD1 and NMATBD2. It was also to confirm that they were accustomed enough to the driving simulator as they practice before starting to perform the first run.

These results also confirmed that for physiological and performance measurements, the order effect did not exist. However, care has to be taken when interpreting the subjective measurement.

4.4.2 Regarding the effect of Non-Driving Task to the secondary tasks

As explained in the previous chapter, NDT had been set before the driving conditions to test on how the MAT alone gave an effect to the measurements. Figure 4.7 (a) and (b) show the results of Normalized HR and Normalized HRV. A repeated-measures ANOVA within-subject factors were performed for task conditions. There were significant differences of Normalized HR between the conditions, F(1,20)=5.707, p=.002. Post hoc tests were conducted using pairwise comparisons with Bonferroni Method, and the results revealed that there were significant differences between MAT1NDT and MAT2HC (p=0.045), between MAT1NHC and MAT2HC (p=0.002) and between MAT1HC and MAT2HC (p=0.009).

While for the Normalized HRV, the significant differences were found between the conditions, F(1,20)=3.929, p=.012. Post hoc tests were conducted using pairwise comparisons with Bonferroni Method revealed that the differences were significant between MAT2NHC and MAT1HC (p=0.048) and between MAT1HC and MAT2HC (p=0.017).



(a) (b)

Figure 4.7 Mean value of Normalized HR and Normalized HRV

From the result of Normalized HR, the differences only significant when compared with the most difficult task of MAT2HC where they were performing MAT2 under HC. It was also known that, although participants were performing only the MAT without a driving task, the data of Normalized HR and HRV show that there was no difference between performing the task while driving or not. This result shows that, the differences between driving and not driving was found only when performing an easy level of MAT. When performing MAT2, the different was exist.

4.4.3 Physiological Measurements

For physiological measurement, a repeated-measures ANOVA was conducted from the individual means of data. A 2x2x3 mixed factorial design was used, with two sources: Finger BVP and ECG were regarded as between-subjects factors and two independent variables: type of traffic conditions (NHC and HC) and type of secondary Task (NMAT, MAT1, and MAT2) was regarded as within-subjects factors.

4.4.3.1 Mean HR and Normalized HR

Table 4.3 Univariate test results for HR and Normalized HR. MAT= Type of a secondary task, NMAT, MAT1 and MAT2, Sources= Sources of data, BVP and ECG, TC: Traffic conditions, NHC and HC. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when the Mauchly's test showed a violation of sphericity.

Effect	HR				Normalized HR			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**2,47**	**23.571**	**<0.001**	**0.329**	**2,47**	**52.614**	**<0.001**	**0.691**
Sources (S)	1,48	0.002	0.962	0.000	1,48	0.031	0.860	0.001
MAT x S	2,47	0.417	0.661	0.017	2,47	0.456	0.637	0.019
Traffic Condition (TC)	1,48	0.415	0.522	0.009	**1,48**	**5.595**	**0.022**	**0.104**
TC x S	1,48	0.045	0.834	0.001	1,48	0.312	0.579	0.006
MAT x TC	**1.5,74.2**	**39.649**	**<0.001**	**0.452**	**1.8, 85.6**	**5.259**	**0.009**	**0.099**
MAT x TC x S	1.5,74.2	0.033	0.937	0.001	1.8, 85.6	0.017	0.976	0.000

Table 4.3 shows a univariate test on HR and Normalized HRV. As we can see on the table, the significant main effect was found on the secondary tasks. Significant main effects of Traffic Condition (TC) also found on the Normalized HR F (1,48) =5. 595, p=0. 022, η_p^2 =0. 104. The interaction between MAT and Traffic conditions also found on both HR and Normalized HR.

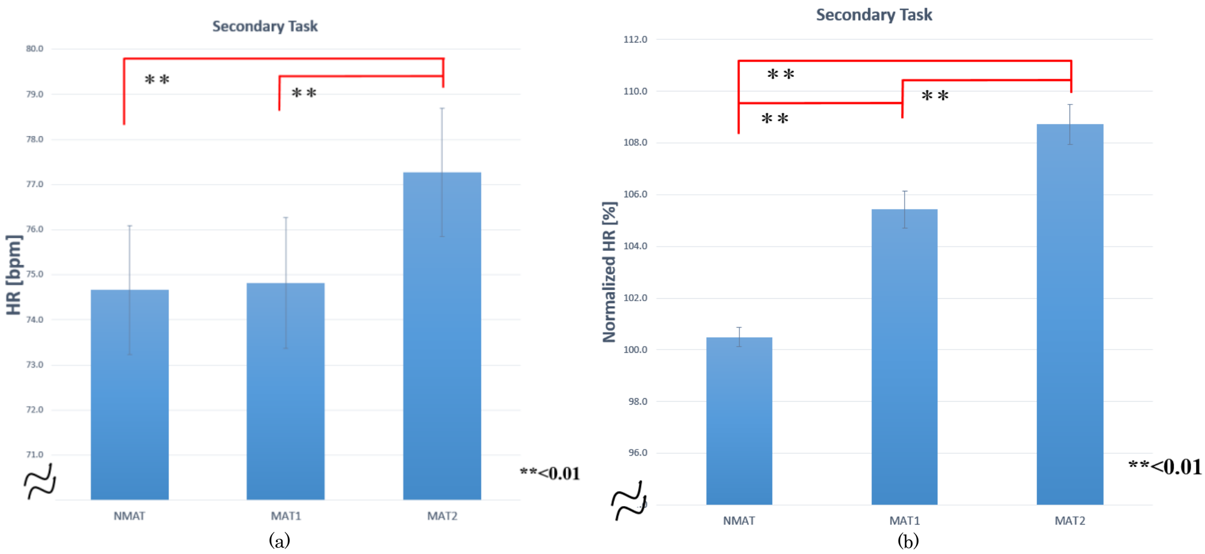


Figure 4.8 Mean value of HR and Normalized HR according to type of secondary Tasks

Figure 4.8 shows a mean value of HR and Normalized HR according to the type of secondary tasks. From the graph, HR was the highest under MAT2 condition (77.27) and the lowest value found under NMAT condition (74.66). While for Normalized HR, the highest value also found on MAT2 (108.72%) while NMAT recorded the lowest (100.50%). A posthoc test using pairwise comparisons with Bonferroni Method revealed that significant differences were found on between NMAT-MAT1 (p<0.001) and NMAT-MAT2 (p<0.001) for HR. The difference between NMAT and MAT1 was not identified in HR. While for Normalized HR, significant differences were found between these three types of secondary tasks as shown in Figure 4.8 (b).

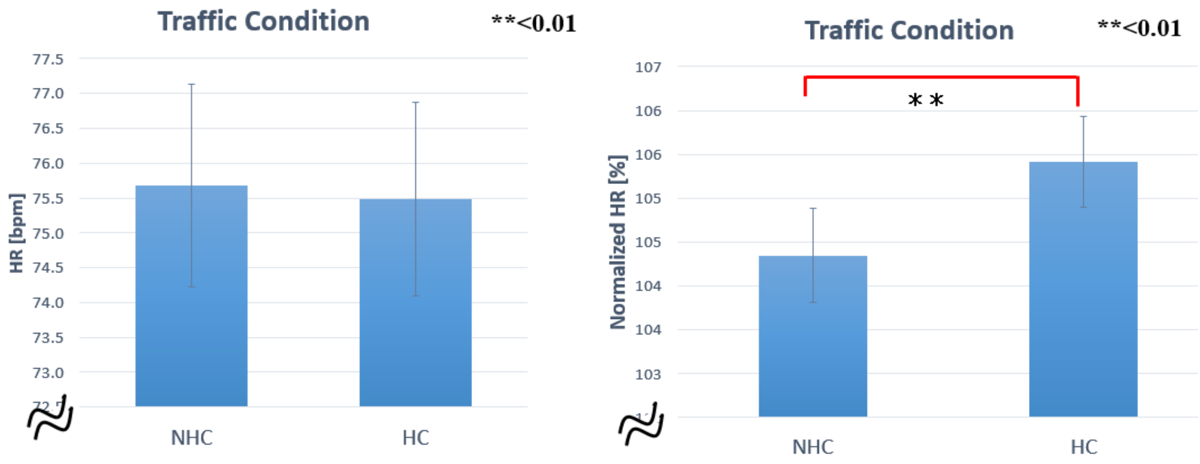


Figure 4.9: Mean value of HR and Normalized HR according to the type of traffic conditions

Figure 4.9 shows a mean value of HR and Normalized HR according to the type of traffic conditions. A significant difference between NHC and HC was found in Normalized HR (p=0. 022). The value of Normalized HR was higher during hazardous condition (105.41%) compared to Non-Hazardous condition (104.35%).

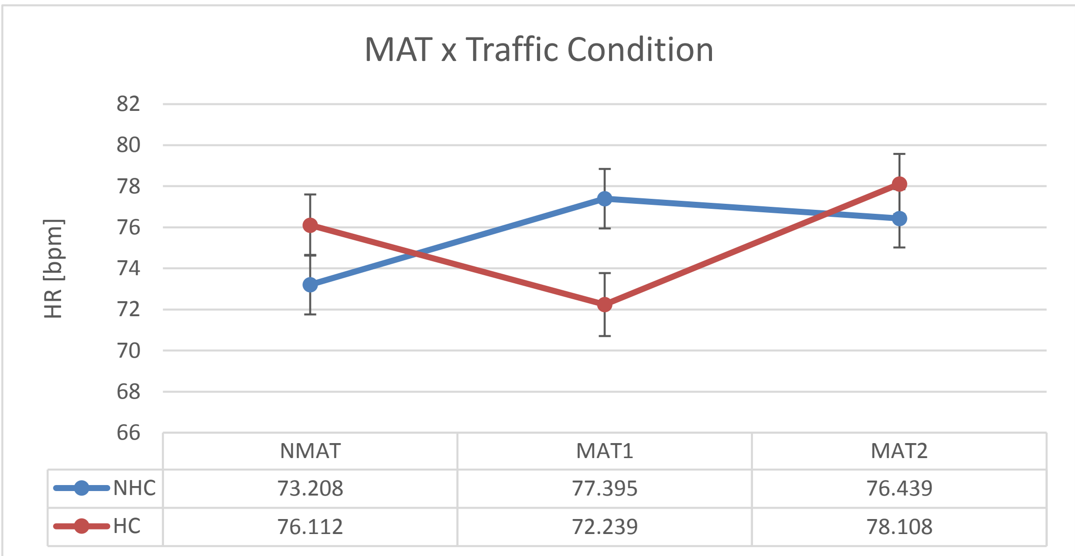


Figure 4.10 Interaction between MAT and Traffic Condition for HR

As we can see from Table 4.3, the interaction between MAT and Traffic conditions was significant for both HR and Normalized HR. Figure 4.10 shows a graph of interaction between MAT and Traffic Conditions on HR. From the graph, under NHC traffic condition, HR was increased when participants accomplishing the MAT1 and slightly dropped when accomplishing a difficult level of MAT. While under HC, HR was significantly dropped to the lowest when accomplishing MAT1 and increased drastically when accomplishing MAT2. This phenomenon can be explained as the uncertainty in predicting the HR value under hazardous conditions.

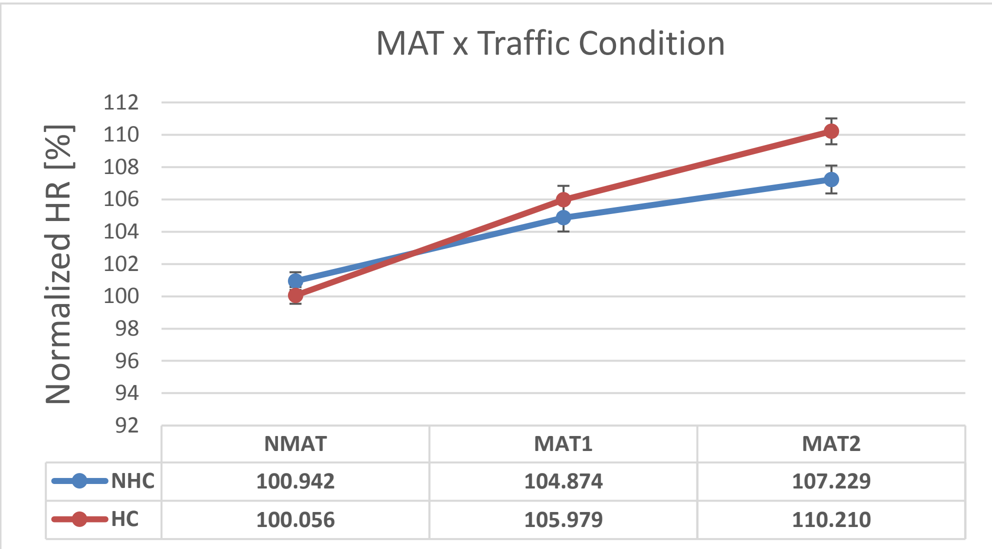


Figure 4.11 Interaction between MAT and Traffic conditions for Normalized HR

Figure 4.11 shows the interaction graph between MAT and Traffic conditions for Normalized HR. In this data, a clear trend could be seen from the figure which is the increasing level of difficulties of MA Task could lead to increasing value of Normalized HR under both traffic conditions. This result could be interpreted as raising the level of difficulties of either MAT task or level of complexity on traffic conditions will lead to further increased on Normalized HR data.

From the value of effect size data, we can see that the interaction between MAT and Traffic condition gives the biggest effect on the value of HR, η_p^2 =0.452. While for Normalized HR, MAT gives higher effect to the value with η_p^2 =0.691.

4.4.3.2 Mean HRV and Normalized HRV

Table 4.4 Univariate test results for HRV and Normalized HRV. MAT= Type of a secondary task, NMAT, MAT1 and MAT2, Sources= Sources of data, BVP and ECG, TC: Traffic conditions, NHC and HC. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when the Mauchly's test showed a violation of sphericity.

Effect	HRV				Normalized HRV			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	2,47	2.253	0.116	0.087	2,47	2.265	0.115	0.088
Sources (S)	1,48	0.502	0.482	0.010	1,48	0.114	0.737	0.002
MAT x S	2,47	0.232	0.794	0.010	**2,47**	**4.262**	**0.020**	**0.154**
Traffic Condition (TC)	1,48	0.026	0.872	0.001	1,48	0.522	0.474	0.011
TC x S	1,48	0.080	0.778	0.002	1,48	0.807	0.374	0.017
MAT x TC	**1.5,73.5**	**10.517**	**<0.000**	**0.180**	1.6, 76.8	3.115	0.061	0.061
MAT x TC x S	1.5,73.5	0.195	0.764	0.004	1.6, 76.8	1.840	0.173	0.037

Table 4.4 shows univariate test results for HRV and Normalized HRV data. As shown in the table, significant main effects were found in the interaction between MAT and Traffic conditions for HRV value and interaction between MAT and Sources of data for Normalized HRV value.

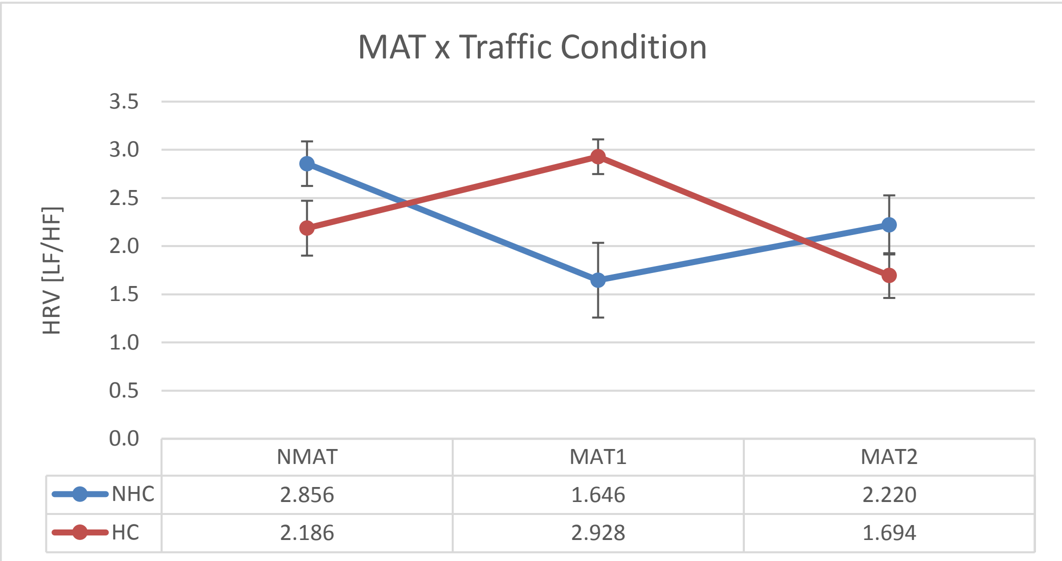


Figure 4.12 Interaction between MAT and Traffic conditions for HRV

Figure 4.12 shows a graph of interaction between MAT and Traffic conditions for HRV data. As shown in the graph, the value of HRV was different according to the traffic conditions. The main difference was found when the participants were accomplishing MAT1. Under NHC the value was dropped to the lowest on MAT1 and increased again when accomplishing the MAT2.

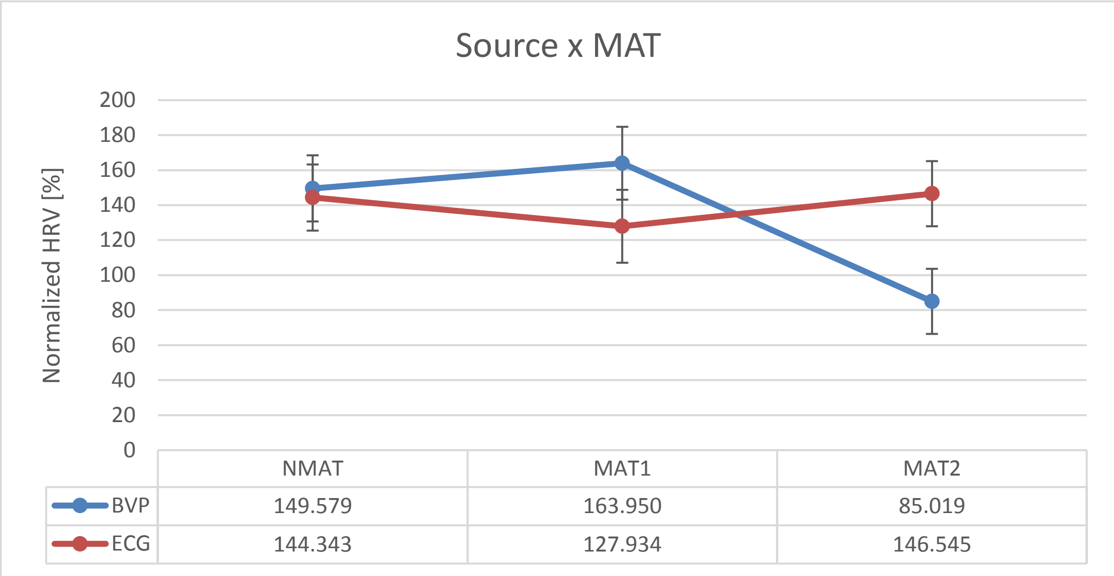


Figure 4.13 Interaction between Source of data and MAT on Normalized HRV

The interaction between sources of data and MAT was found significant for Normalized HRV data, F (2,47) =4. 262, p=0. 020, η_p^2 =0. 154. Figure 4.13 shows a graph of interaction between the source of data from and MAT on Normalized HRV. From the graph, the data from BVP was dropped when participants were accomplishing the difficult level of MAT. As our discussion regarding HRV, the more difficult the task is, HRV will decrease, based on this data, BVP is more sensitive to the changes of level of difficulties compared to data from the ECG.

4.4.4 Comparison between BVP data and ECG data

As one of our interests for this experiment was to evaluate the usage of BVP sensor as a tool to assess mental workload, we compare the data with ECG data. The usage of BVP is practical to be used in real driving conditions as it is non-intrusive to drivers. For this evaluation, we only focus on HR data from BVP and ECG.

To evaluate the usage of BVP as a tool to measure mental workload, a paired-samples t-test was conducted to compare HR data from BVP and ECG in driving conditions. The differences were not significant in every condition compared to data from the ECG and BVP. HR BVP (M=104. 41, SD=6. 98) and HRV ECG (M=104. 57, SD=6. 70); t (249) =-0.630, p = 0.529.These results suggest that the data from BVP and ECG was the same.

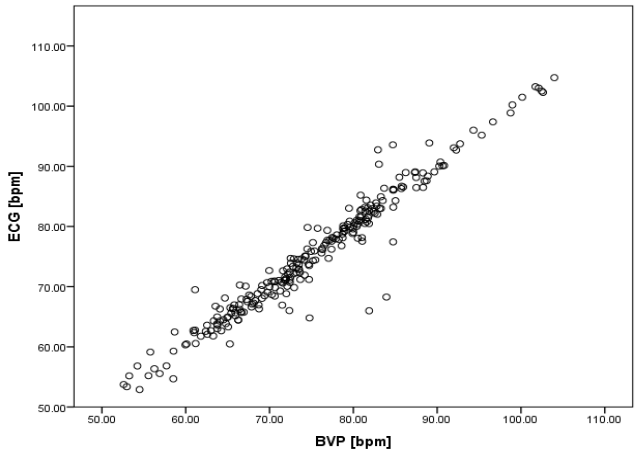


Figure 4.14 Scatterplot of ECG and BVP for HR data

A Pearson product-moment correlation coefficient was computed to assess the relationship between HR from BVP and ECG. There was a positive correlation between the two variables, r=0. 971, n=250, p<0.0005. A scatter plot summaries the results (Figure 4.14). Overall, there was a strong positive correlation between HR from BVP and ECG. Increases in HR from ECG according to driving conditions were correlated with increases in HR from BVP.

To further evaluate the usage of BVP, a repeated-measures ANOVA was computed from the individual means of HR from BVP and ECG data in six driving conditions (NMATNHC, MAT1NHC, MAT2NHC, NMATHC, MAT1HC, and MAT2HC).

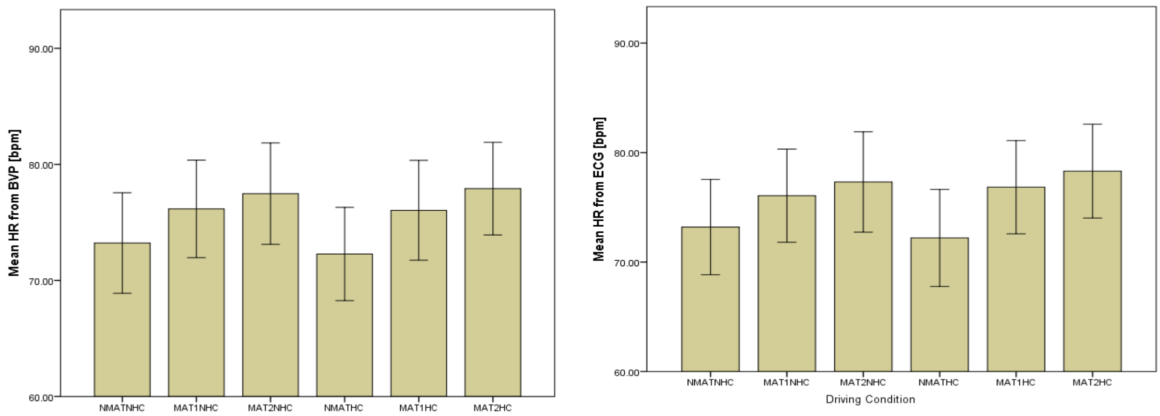


Figure 4.15 Mean Value of HR from ECG and BVP

The average value of HR data from BVP for the MAT intervals was shown in Figure 4.15 (a). As illustrated in the graphs, the highest value was found on MAT2HC (M: 77.91: SD:9.67), and NMATHC was the lowest of HR value with (M: 72.28, SD: 9.72). The results of ANOVA determined that mean HR from BVP differed statistically significantly between driving conditions (F (5, 20) =6. 837, p=0. 001, η_p^2 =0. 631).

While for ECG, as shown in Figure 4.15 (b) the highest value was found on MAT2HC (M: 78.31, SD: 10.40) and NMATHC was the lowest of HR value with (M: 72.20, SD: 10.74). The results of repeated-measures ANOVA reveals that mean HR from ECG differed statistically significantly between driving conditions ($F (5,20) = 7.005$, $p = 0.001$, $\eta_p^2 = 0.637$).

A posthoc test using Bonferroni method was conducted and the results of significant difference between conditions shown in Table 4.5.

Table 4.5 Pairwise Comparisons for HR data from BVP (a) and ECG (b), Significant effects (p<0.05) are shown in bold.

BVP	NMATNHC	MAT1NHC	MAT2NHC	NMATHC	MAT1HC	MAT2HC
NMATNHC	X	.062	**.013**	1.000	.054	**.002**
MAT1NHC	.062	X	.991	**.006**	1.000	.148
MAT2NHC	**.013**	.991	X	**.000**	1.000	1.000
NMATHC	1.000	**.006**	**.000**	X	**.002**	**.000**
MAT1HC	.054	1.000	1.000	**.002**	X	.252
MAT2HC	**.002**	.148	1.000	**.000**	.252	X

ECG	NMATNHC	MAT1NHC	MAT2NHC	NMATHC	MAT1HC	MAT2HC
NMATNHC	X	.092	**.003**	1.000	**.009**	**.002**
MAT1NHC	.092	X	1.000	**.019**	1.000	.137
MAT2NHC	**.003**	1.000	X	**.000**	1.000	1.000
NMATHC	1.000	**.019**	**.000**	X	**.000**	**.000**
MAT1HC	**.009**	1.000	1.000	**.000**	X	1.000
MAT2HC	**.002**	.137	1.000	**.000**	1.000	X

When looking at the significant difference between driving conditions, ECG data were able to distinguish more compared to BVP. Especially the difference between NMATNHC and MAT1HC which BVP failed to have a significant difference between these two driving conditions ($p = 0.054$).

For the conclusion of the difference between BVP and ECG, BVP can be used as a tool to evaluate mental workload as it comes from the same source of data. However, when we look at the accuracy to distinguish between driving conditions, ECG data gives more accurate data compared to BVP.

4.4.5 Performance Measurements

4.4.5.1 **Primary task performance**

For the primary task performance, two ways of lateral performance were measured, Standard Deviation of Lateral Position (SDLP) and Steering Entropy [52].

Table 4.5 Univariate test results for SDLP and Normalized SDLP. MAT= Type of a secondary task, NMAT, MAT1 and MAT2, TC: Traffic conditions, NHC and HC. Significant effects ($p<0.05$) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when the Mauchly's test showed a violation of sphericity.

Effect	SDLP				Normalized SDLP			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**2,23**	**24.370**	**<0.001**	**0.679**	**2,23**	**17.717**	**<0.001**	**0.606**
Traffic Condition (TC)	**1,24**	**12.628**	**0.002**	**0.345**	1,24	0.521	0.477	0.021
MAT x TC	1.5, 36.3	1.566	0.219	0.061	2,23	0.436	0.652	0.037

Table 4.5 lists the results of the univariate test for SDLP and Normalized SDLP. From the table, significant main effects of MAT were found on both SDLP and Normalized SDLP. The effect of traffic conditions was also significant for the SDLP. The results of pairwise comparisons and significant differences were shown in Figure 4.16 (a~d).

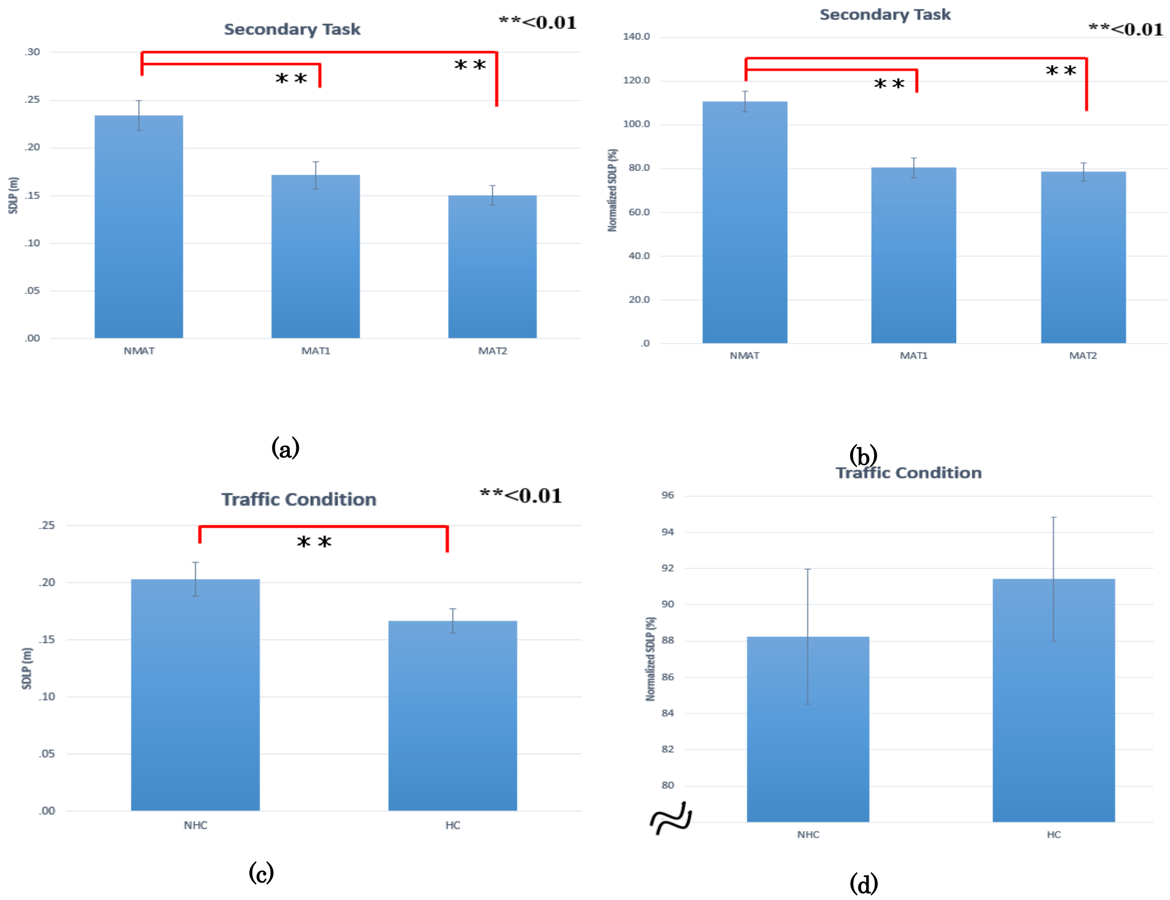


Figure 4.16: Average SDLP and Normalized SDLP according to the type of secondary task.

Figure 4.16(a) and (b) depict a mean value of the SDLP and Normalized SDLP according to the type of secondary tasks. Participants show less lane variability when driving under Mathematical Arithmetic Tasks. Post hoc tests were performed, and the results revealed that there were significant differences between NMAT and MAT1 (p=. 000) and between NMAT and MAT2 (p=. 000). There was no significant difference found between MAT1 and MAT2 (p=1. 000). From the graph, SDLP was below the value of 100% when participants were opposed to Mathematical Arithmetic Task either MAT1 or MAT2.

Figure 4.16 (c) shows the mean value of measurements according to traffic condition. From the results, participants swerve less when driving on HC traffic conditions. This result also different with our expectations that the driver will swerve more under the hazardous conditions.

Table 4.6 Univariate test results for Steering Entropy and Normalized SE. MAT= Type of a secondary task, NMAT, MAT1 and MAT2, TC: Traffic conditions, NHC and HC. Significant effects (p<0.05) are shown in bold.

Effect	Steering Entropy				Normalized SE			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**2,23**	**13.384**	**<0.001**	**0.538**	**2,23**	**12.183**	**<0.001**	**0.514**
Traffic Condition (TC)	1,24	0.719	0.405	0.029	1,24	0.609	0.443	0.025
MAT x TC	2,23	0.228	0.798	0.019	2,23	0.025	0.976	0.002

Table 4.6s shows the results of the univariate test for Steering Entropy and Normalized steering entropy (SE). From the table, significant main effects of MAT were found on both SE and Normalized SE. The results of pairwise comparisons and significant differences were shown in Figure 4.17 (a~d).

Figure 4.17 (a) and (b) shows a mean value of Steering Entropy and Normalized Steering Entropy as functions of a secondary task. Participants show higher Steering Entropy value when driving under Mathematical Arithmetic Tasks. Post hoc tests were conducted using pairwise comparisons with Bonferroni Method, and the results revealed that the differences were significant between NMAT and MAT1 (p<. 001) and between NMAT and MAT2 (p<. 001). There was no significant difference found between MAT1 and MAT2 (p=0. 915). There was also no significant main effect of Traffic Condition found on Steering Entropy and Normalized Steering Entropy. There was also no interaction between MAT and Traffic conditions.

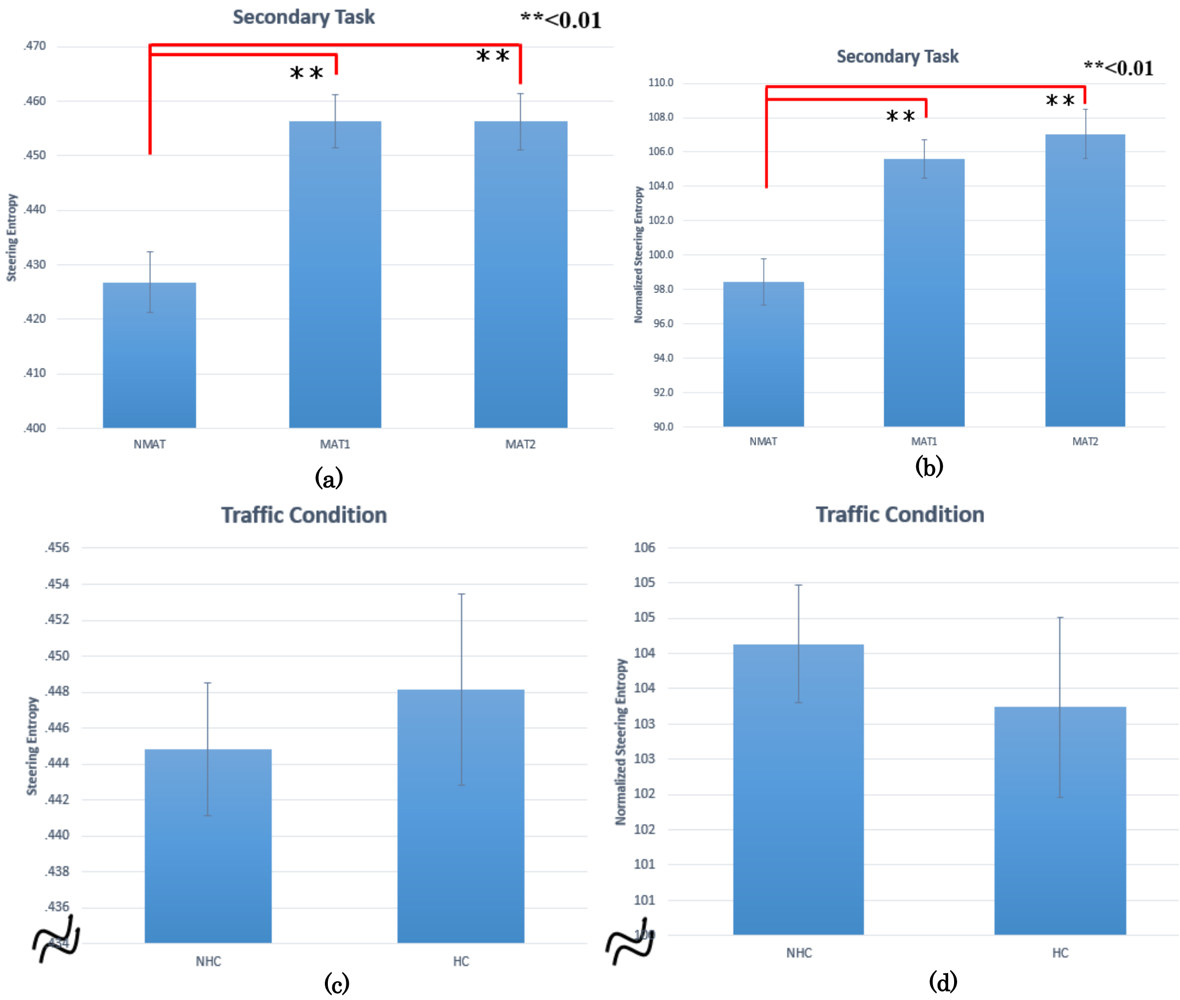


Figure 4.17 Average value of Steering Entropy and Normalized SE

4.4.5.2 Secondary Task performance

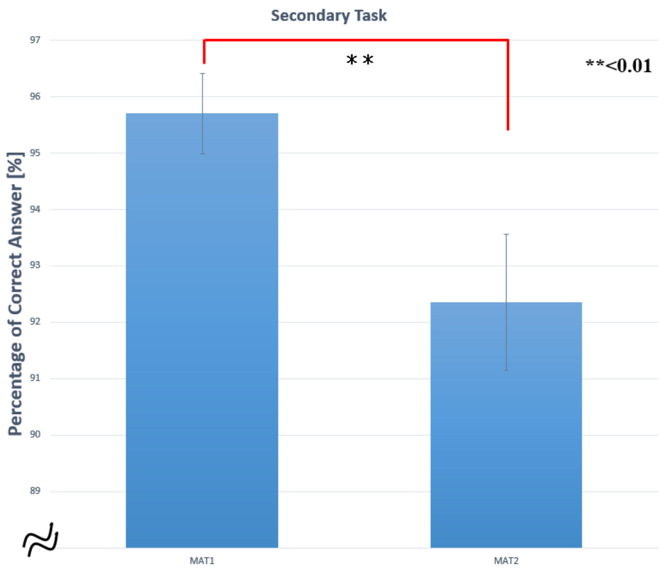


Figure 4.18 Mean value of Percentage of the Correct answer (±SE) for secondary task.

Table 4.7 Univariate test results for the percentage of the correct answer. MAT= Type of a secondary task, NMAT, MAT1 and MAT2, TC: Traffic conditions, NHC and HC. Significant effects (p<0.05) are shown in bold.

Effect	Percentage of Correct Answer			
	df 1,2	F	p	η_p^2
MAT	**1,24**	**10.629**	**0.003**	**0.307**
Traffic Condition (TC)	1,24	2.971	0.098	0.110
MAT x TC	1,24	1.165	0.291	0.046

Figure 4.18 shows a percentage of correct answer of a secondary task. From the graph, the performance of participants dropped when performing MAT2. A repeated-measures ANOVA was conducted, and the results show that a significant main effect of MAT found in the percentage of the correct answer, F (1,24) =10. 629, p=. 003, η_p^2 =0. 307, as shown in Table 4.7. Post hoc tests were conducted using pairwise comparisons with Bonferroni Method, and the result revealed that there wa.a significant difference. Between MAT1 and MAT2 (p=0. 001). The significant main effect was not found for Traffic Condition on the percentage of the correct answer, F (1,24) =2. 971, p=. 098, η_p^2, =0. 110. The interaction between MAT and TC also not significant, F (1,24) =1. 165, p=0. 291, η_p^2 =0. 046.

From the results, the performance of secondary task was dropped when they are accomplishing a difficult level of secondary tasks.

4.4.6 Subjective measurements

For subjective measurements, two kinds of subjective measurements were analyzed, the NASA-TLX and RSME.

Table 4.8 Univariate test results for NASA-TLX and RSME. MAT= Type of a secondary task, NMAT, MAT1 and MAT2, TC: Traffic conditions, NHC and HC. Significant effects (p<0.05) are shown in bold.

Effect	NASA-TLX				RSME			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**1.6, 37.4**	**74.316**	**<0.001**	**0.756**	**2,23**	**64.266**	**<0.001**	**0.848**
Traffic Condition (TC)	**1,24**	**33.289**	**<0.001**	**0.581**	**1,24**	**33.285**	**<0.001**	**0.581**
MAT x TC	2,23	2.285	0.124	0.166	**2,23**	**6.965**	**0.004**	**0.377**

Table 4.8 shows univariate test results for subjective measurements of NASA-TLX and RSME. From the table, significant main effects were found for MAT and Traffic conditions for both NASA-TLX and RSME. The interaction between MAT and Traffic conditions was found significant for RSME.

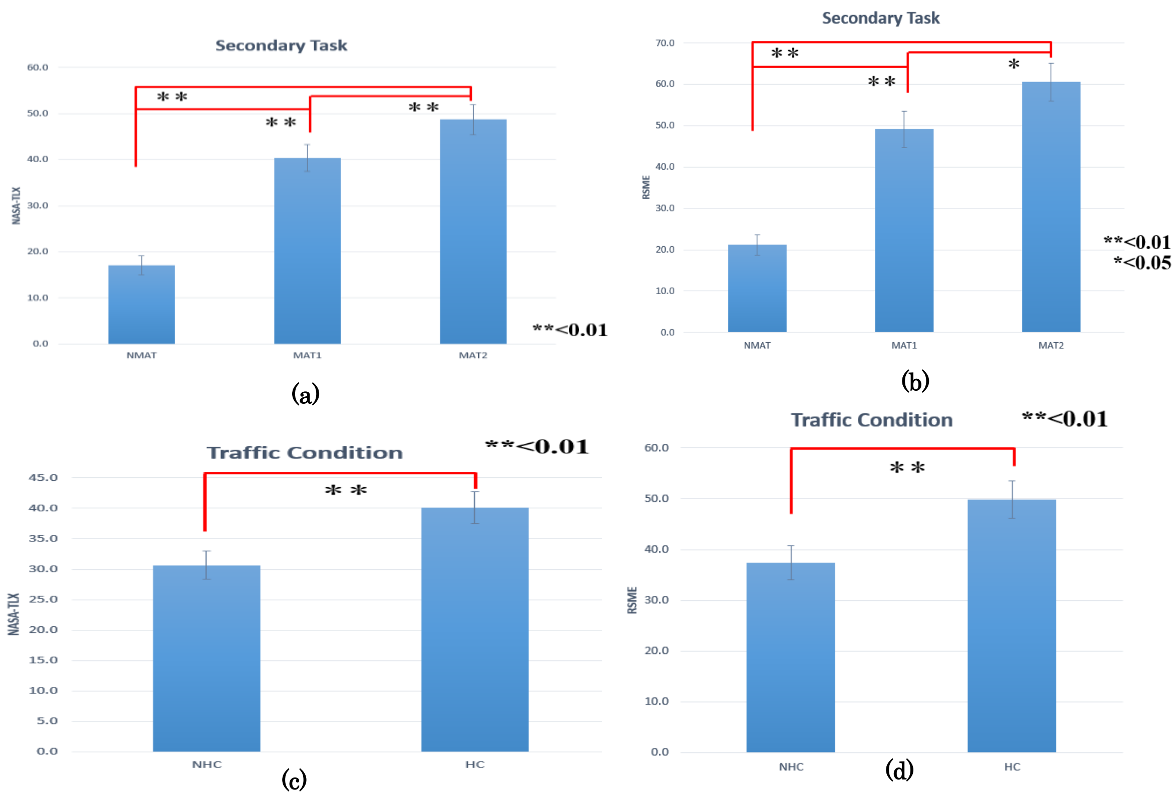


Figure 4.19 Mean value of NASA-TLX (±SE) and RSME (±SE)

Figure 4.19 (a)(b) show the rating for NASA-TLX and RSME according to the secondary tasks. Post hoc tests were conducted using pairwise comparisons with Bonferroni Method, and the results revealed that the differences were significance between NMAT and MAT1 (p=0.000), between NMAT and MAT2 (p=0.000) and between MAT1 and MAT2 (p=0.003) for NASA-TLX. For RSME, the significant differences were found between NMAT and MAT1 (p=0. 000), between NMAT and MAT2 (p=0. 000) and between MAT1 and MAT2 (p=0. 013) respectively. Both of the measurements indicate that MAT2 gives the highest value among others. NMAT recorded as the lowest among others.

Figure 4.19 (c~d) shows graphs of the mean value for NASA-TLX and RSME according to traffic conditions. The main effect of Traffic Condition was also significant F(1,24)=33.285, p =0.000. Post hoc tests revealed that there were significant differences between NMAT and MAT1 (p=0.000), between NMAT and MAT2 (p=0.000) and between MAT1 and MAT2 (p=0.013). Participants rated driving in hazardous conditions a higher mental workload, then non-hazardous condition.

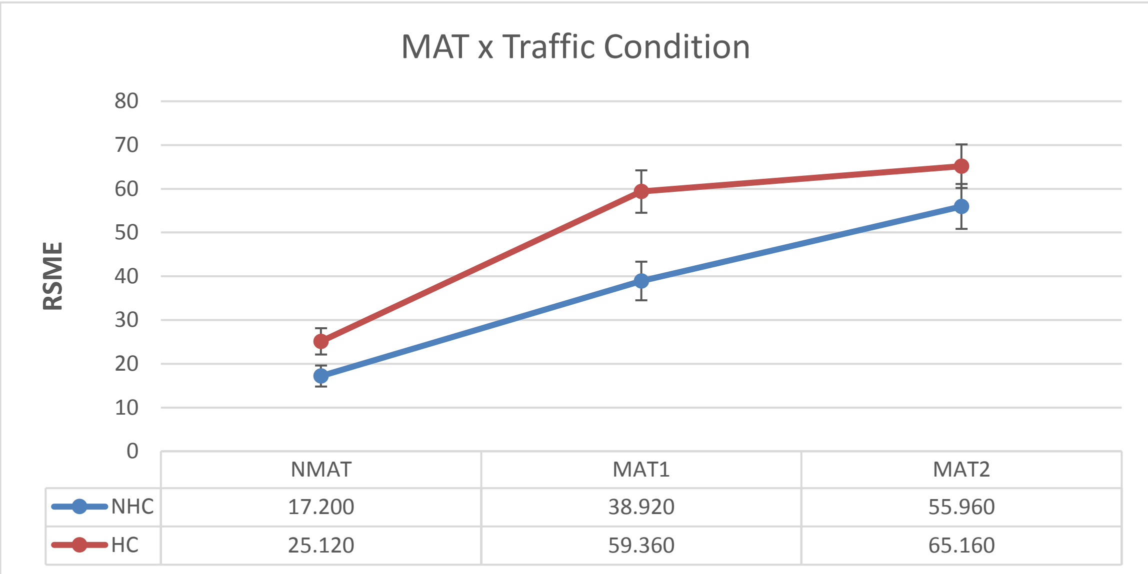


Figure 4.20 Interaction between MAT and Traffic Condition for RSME

Figure 4.20 shows the interaction between MAT and Traffic conditions for RSME. According to the graph, the increasing level of difficulties from secondary task and under more complex traffic condition will give a higher value of mental workload to the drivers. Both of the results from subjective measurements suggest that participants rated the most difficult task were during MAT2HC when they were performing a difficult level of the secondary task under hazardous condition.

4.5 Discussion

4.5.1 Comparison between BVP and ECG data

Evaluation of usage of BVP to assess mental workload under different driving conditions. Results from the HR correlate highly between BVP and ECG. Further analysis of the differences reveals that, although the data from BVP was the same with ECG, regarding the accuracy of distinguishing between driving conditions, ECG data was more accurate. This result is maybe due to the nature of the data which the data from ECG is taken from direct sources of heart rate, while the data from BVP is taken from the blood circulation on the finger. Overall, we think that BVP also can be used as mental workload measurements as it is non-intrusive to the drivers and the data was robust.

4.5.2 The main effects of driving conditions on the measurements.

In the introduction, we explained that we would like to explore the internal and external factors of mental workload by the manipulating levels of difficulties of tasks. In this experiment, MAT was set as an internal factor of mental workload by manipulating the level of difficulties of secondary tasks (NMAT-MAT1-MAT2).

While traffic conditions of NHC and HC were regarded as external factors. By manipulating the difficulties of traffic conditions with the assumption that while driving on HC situation, participants had to pay a proper attention continuously and controlling the vehicle safely so that they not meet with accidents. In this traffic condition, the demand of mental workload was increased.

Table 4.9 Summary of Significant main effects of driving conditions on the measurements

		MAT (p)	η_p^2	Traffic Condition (p)	η_p^2	Traffic Condition × MAT (p)	η_p^2
Physiological	HR	<0.001	0.329	0.522	0.009	<0.001	0.452
	Normalized HR	<0.001	0.691	0.022	0.104	0.009	0.099
	HRV	0.116	0.087	0.872	0.001	<0.001	0.180
	Normalized HRV	0.115	0.088	0.474	0.011	0.061	0.061
Performance	SDLP	<0.001	0.679	0.002	0.345	0.219	0.061
	Normalized SDLP	<0.001	0.606	0.477	0.021	0.652	0.037
	SE	<0.001	0.538	0.405	0.029	0.798	0.019
	Normalized SE	<0.001	0.514	0.443	0.025	0.976	0.002
	Percentage of Correct answer	0.003	0.307	0.098	0.110	0.291	0.046
Subjective	NASA-TLX	<0.001	0.756	<0.001	0.581	0.124	0.166
	RSME	<0.001	0.848	<0.001	0.581	0.004	0.377

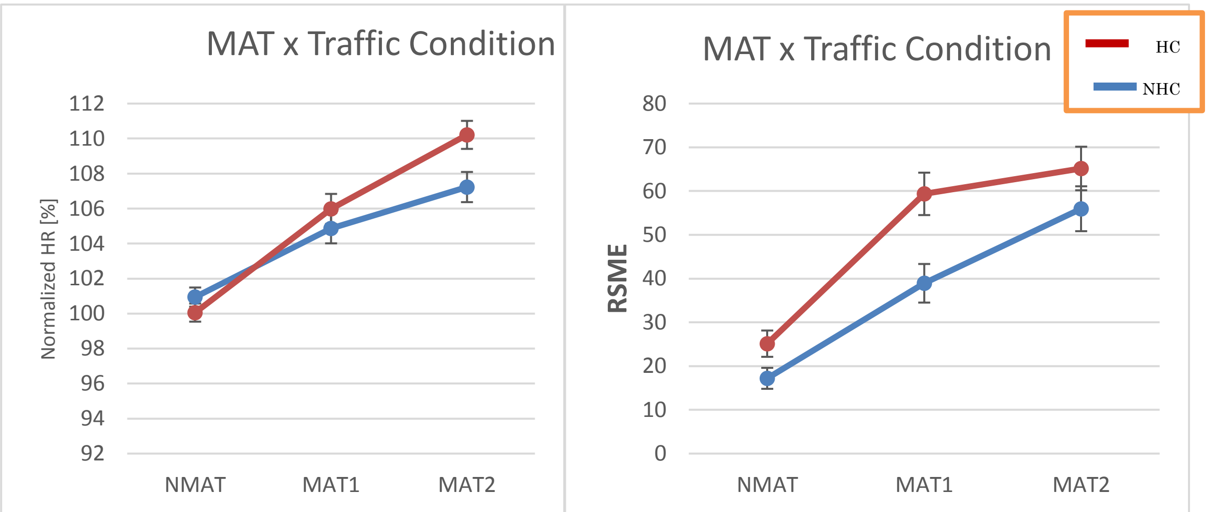
Table 4.9 shows a summary result of significant main effects on driving conditions based each measurement groups. A yellow color of the column represents a significant main effect based on p-value less than 0.5. While the blue color of the column indicates the value of the effect size of partial eta square which was dominant among other effects (MAT, Traffic Condition, and interaction between Traffic condition and MAT) in the measurement.

Based on the results, significant main effects of MAT were found in almost all measurements except for HRV and normalized HRV. While for traffic conditions, a significant effect was found in Normalized HR, SDLP, and both subjective measurements. Based on the effect size of partial eta square, it can be concluded that for this experiment, the effect of secondary task was bigger than traffic conditions. It could be interpreted as, in this experiment, internal factors affect more than external factors of mental workload.

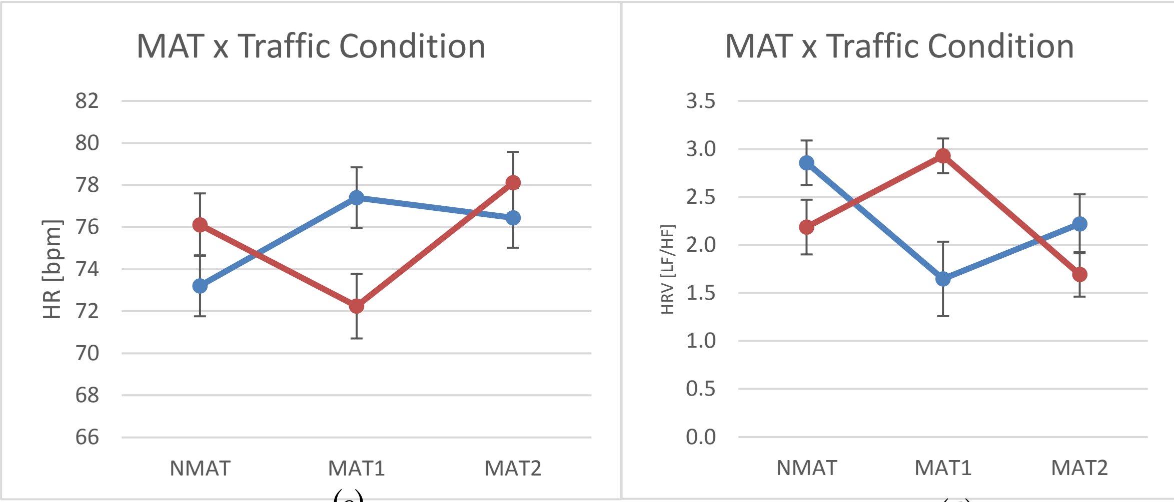
One of the interpretations for this result is while performing the tasks, because of demand from secondary task was higher than a primary task, participants may regard the secondary task as the main task.

4.5.3 Interaction between internal and external factors of mental workload

To discuss the interaction between secondary task and traffic conditions, we refer to Table 4.9. Significant interactions between MAT and Traffic Conditions were found on HR, Normalized HR and HRV data for physiological and RSME for subjective measurement.



(a) (b)



(c) (d)

Figure 4.21 (a~d) Interaction between MAT and Traffic Condition

Figure 4.21 shows graphs of interaction between MAT and Traffic Condition that we found significant. A similar trend we can see for Normalized HR and RSME where the higher the level of difficulties of the task the higher indication of mental workload measurements. This indication can be interpreted as the more difficult the task from internal and external factors, resulted in the higher mental workload.

However, a different interaction between secondary task and traffic conditions for HR and HRV data. For HR data, in HC condition, while accomplishing the easy level of secondary task, the HR data shows a decreased. It was risen again in MAT2 conditions to be the highest. The reverse pattern could be seen on HRV, as the common idea of the higher the mental workload is HR will increase while HRV decrease.

To explain this phenomenon, we think that the trade-off between the difficulty in traffic conditions and increasing difficulties on internal factors of mental workload. This trade-off maybe could be explained by looking at the single resource theory by Kahneman [21]. It also depends on the allocation policy in a human operator.

When we look at the overall results and only consider the results from normalization data as it is more accurate to estimate the workload data, we can conclude that the combination of internal and external factors will further increase the mental workload.

4.5.4 Relationship between Mental Workload and Driving Performance

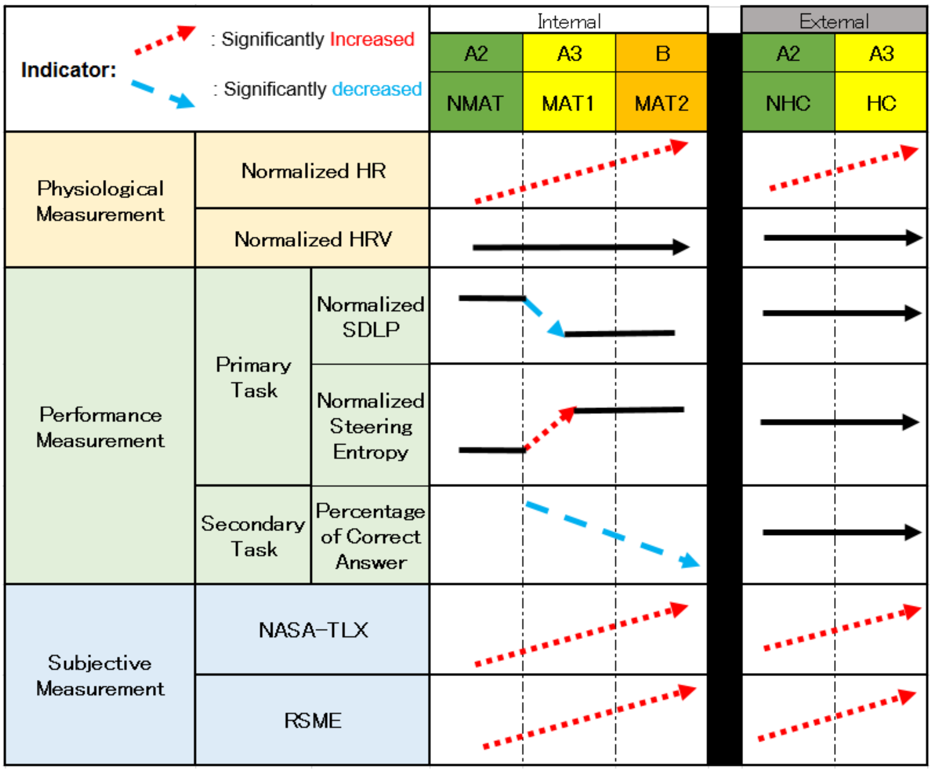


Figure 4.22 Overall results for evaluation of mental workload

Figure 4.22 shows the overall results for evaluation of mental workload. For internal factors, all measurement data show the increasing trend of mental workload except for SDLP shows an improvement of lateral control performance.

When participants were not performing any secondary task under NMAT condition, they were able to maintain the performance, and the subjective measurements show that the mental workload was the lowest. The data from HR and Normalized HR of physiological measurements also indicate that the mental workload was at the lowest level. This condition could be interpreted as they were in Region A2.

While participants were performing the easy level of secondary task (MAT1), the data of subjective measurements started to increase. The data from Normalized HR also indicate a significant increasing compared to the data in NMAT condition. While for main task performance measurements, the data of Normalized SDLP start to decrease and the Normalized Steering Entropy started to increase. However, the data of HRV and Normalized HRV shows no significant difference between performing NMAT and MAT1. We conclude that when opposing MAT1, they were in Region A3. Relationship between driving performance and mental workload

When looking at the overall data in Figure 4.22, we can see that the SDLP and Steering Entropy show changes between NMAT and MAT1. For physiological measurements, HR shows an increase in from NMAT and MAT1 and also from MAT1 to MAT2. The same results of increasing mental workload we can see for subjective measurements. These results can be interpreted as for this experiment the internal factors affect the decrease in performance and increase in mental workload. This finding was contradictory with the idea of the existence of the region A3 in de Waard's model where participants exerted their effort to maintain performance. In the model, because of making their effort to cope with the increasing demand of mental workload, the changes in physiological measurements could be found first before the changes in performance measurements. While in this experiment, it seems that performance and mental workload change concurrently each other in the same region.

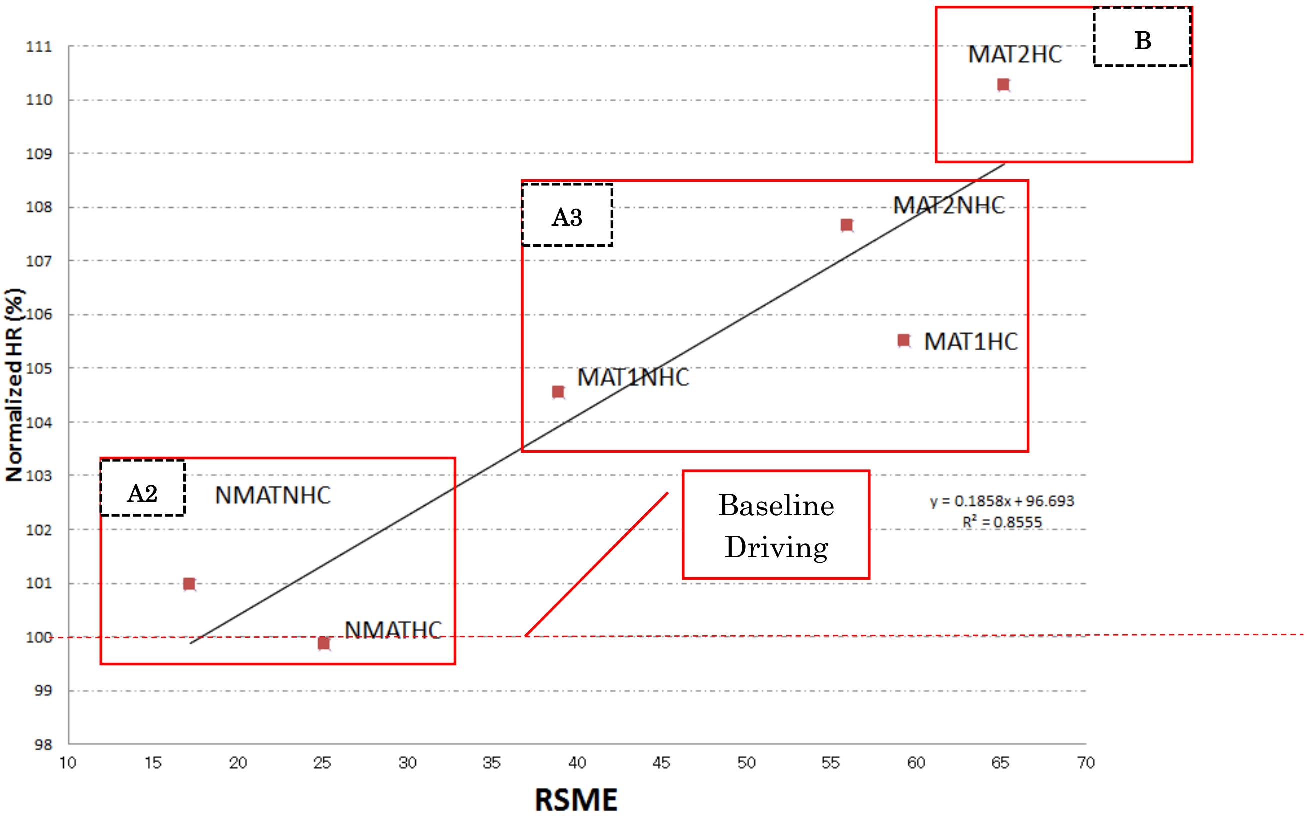


Figure 4.23 Mapping between Normalized HR and RSME

Figure 4.23 shows a mapping graph between a physiological measurement of Normalized HR and a subjective measurement of RSME. As shown in the mapping chart, under NMATNHC and NMATHC conditions where participants have not performed any MAT, Normalized HR was almost 100% and close to the value of Baseline Driving (100%). While for RSME, the participant's rate both NMATNHC and NMATHC were the lowest among others. By considering these points, it is appropriate to map these two conditions under region A2 where the drivers were in low workload.

Under the MAT1NHC the Normalized HR was increased to above 104%, and the rating of RSME was also higher than NMATHC. It means that the demand of workload for the task was greater than performing NMATHC. They were in a different region as before. While in both conditions of MAT2NHC and MAT1HC, the Normalized HR and RSME were about in the same location. It is suitable to map these three conditions into Region A3, where participants were exerting their effort to cope with the increasing demand of workload.

Finally, under the MAT2HC driving condition, the Normalized HR was increased to above 110% and was rated as the most difficult task among others. It is appropriate to map MAT2HC as under region B where the participants were actually in an overload condition.

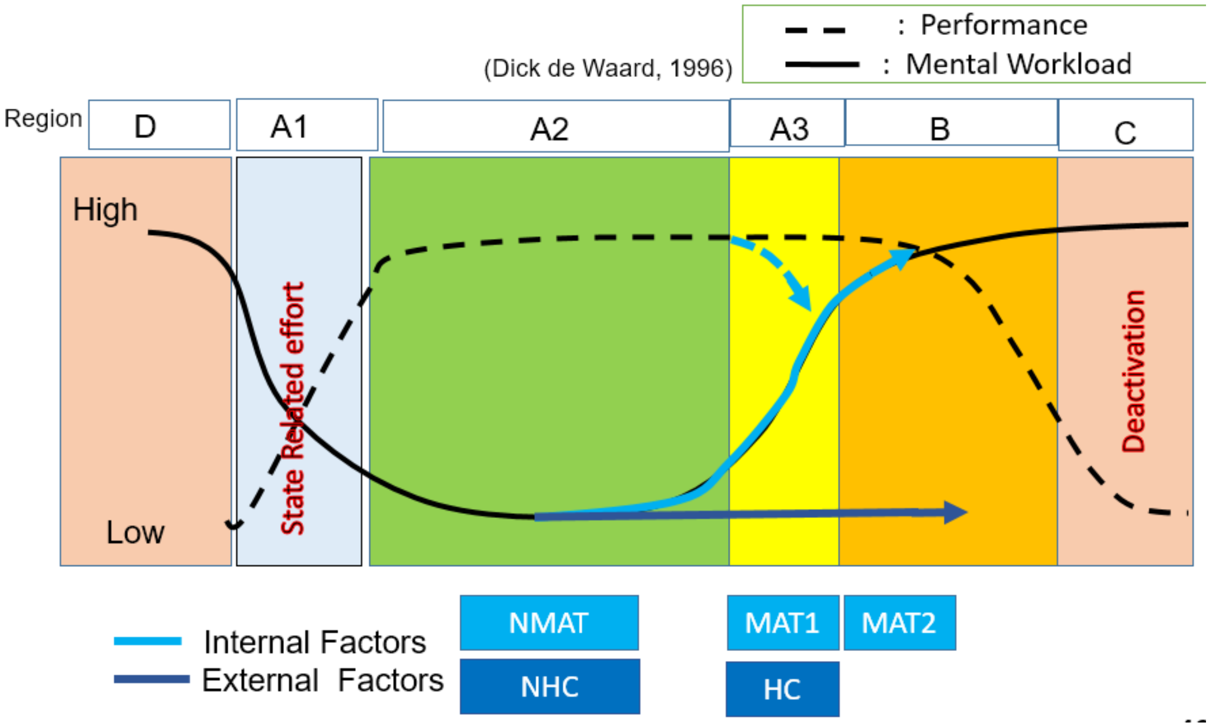


Figure 4.24 The relationship between Performance and Mental Workload

Figure 4.24 shows a relationship between performance and mental workload based on de Waard's 1996. We tried to map the relationship with the data we get from our experiments. From overall results, we can see that for internal factors, the performance starts to change when the participants tried to solve MAT. The data from physiological and subjective measurements show that the mental workload was also increased when they were accomplishing the MAT. We can draw a line as in Figure 4-23 that under internal factors the performance starts to deteriorate as the mental workload increased. We also argued that Region A3 which explained as a task-related effort does not exist when participants, contacting with internal factors in this experiment.

Reversely thinking, for external factors, as we do not find any changes in physiological measurements when contacting with external factors, it may show that there is maybe no effect of external factors in this experiment of mental workload.

Another point to discuss regarding the performance is the reducing of SDLP and increasing Steering Entropy while accomplishing the mental arithmetic task. The reducing SDLP is consistent with our finding for Experiment#1. For further explanation we added another measurement which is Steering Entropy to give an answer whether this phenomenon can regard as improvement of performance or otherwise decrement. Based on the results of this experiment we concluded that the reduced SDLP was resulted from internal factors and does not mean an improvement in driving performance.

In this experiment, there were three experimental conditions that we think influence the results and needed a further research, 1. The primary task was to maintain in the middle of a straight lane with a wide lane width (3.5m wide). 2: The secondary task used in this experiment was only from the auditory task which not required any disturbance on the sight of the participants. 3: Participants were asked to follow a lead vehicle (LV) and have to pay intention on only the movement of LV.

Another possible explanation is that participants were told to drive in the middle of the road and follow the FV, they had no discrete value of performance. However, when exerted to MAT1, they started to put their effort to the steering and less variability lateral position can be seen from the SDLP data and also a higher steering entropy data can be seen from the performance.

These results support our assumption that the decrease in SDLP value in Experiment#1 does not mean the performance of drivers deteriorate, but merely to indicate that they are not paying their attention to the driving and focusing on the internal factors resulted in their swerve less in the SDLP.

4.5.5 Factors of increasing mental workload

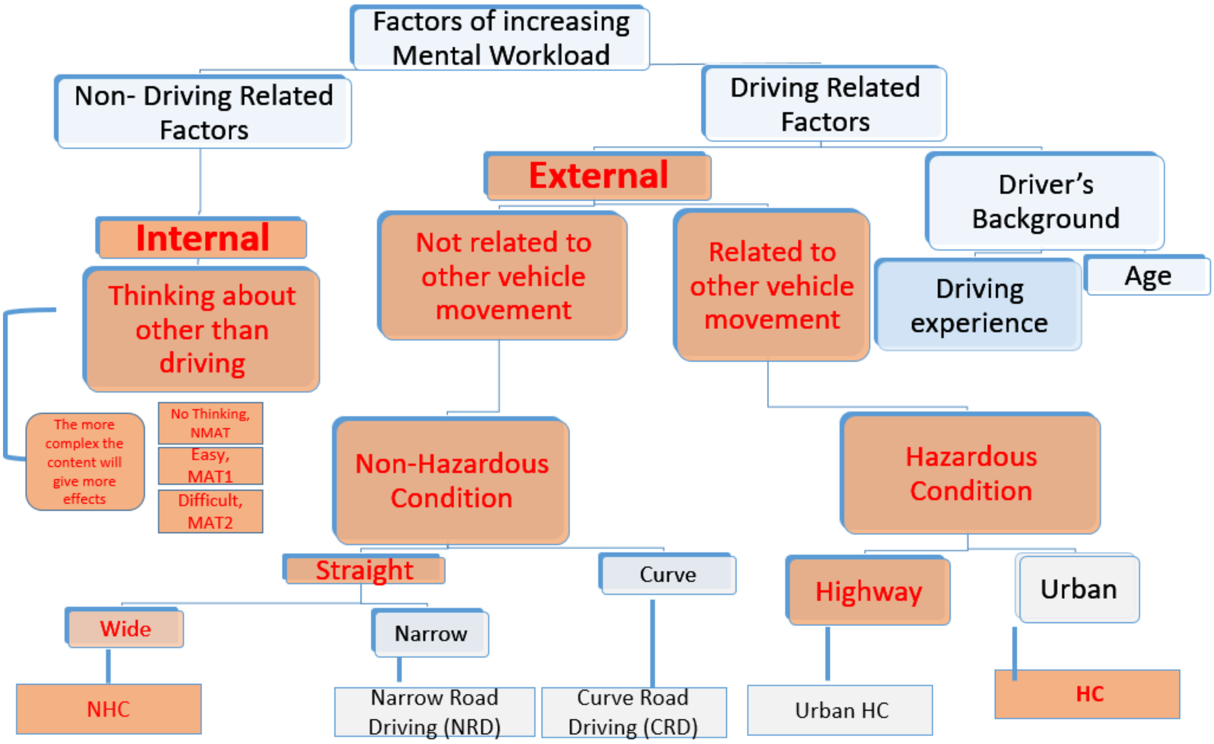


Figure 4.25 Factors of increasing mental workload in Experiment#2

Figure 4.25 shows factors of increasing mental workload that we think to affect the mental workload measurement in Experiment#2. As in this Experiment#2, we are focusing on internal factors, based on the data from physiological and subjective measurements; we were able to distinguish between levels of difficulties in internal factors by designing NMAT-MAT1 and MAT2. It is also appropriate to interprete as the more complex the content that a driver's thinking, the more effect we can observe on physiological measurements.

4.6 Conclusion

Conclusions from this experiment are as follows:

1. MAT that represents the internal factors gives more effects to the participants than Traffic Conditions which represent the external factors. The more complex the content is, the more mental workload.
2. The combination of external and internal will further increase mental workload.
3. Through the normalization of data and combination between physiological and subjective measurement, we were able to map the driving conditions MAT2HC-Region B (Overload), NMATNHC-NMATHC: Region A2 (Low load), MAT1NHC, MAT1HC, MAT2NHC: Region A3 (Task-related effort)
4. SDLP decrease when imposing with MAT and no further decrease when opposing with the more difficult level of MAT. The decreasing value of SDLP does not mean an improvement in driving performance.

Chapter 5 External factors of mental workload

5.1 General overview and aims

In Experiment#2, we found that the main effects between driving conditions were mainly from internal factors. Some of the measurements such as normalized HR show that further increase the difficulties in internal factors will give a higher mental workload to drivers. For performance measurements, the differences only found for between NMAT and MAT1.

As we focused internal factors of mental workload in Experiment#2, and the results show that the traffic conditions did not give effect to the drivers, we designed Experiment#3 to find another factor that affects the driver's mental workload. In Experiment#3, we focused on external factors in increasing mental workload. Further research on a different kind of traffic condition and in a different environment has to be taken to confirm the phenomena especially regarding the lateral control of the vehicle. It is worth finding the relationship between those measurements in a different kind of traffic condition and geometry.

The main objective of this experiment was to investigate the relationship between increased workload and driving performance by manipulating levels of difficulties of the external factors of mental workload. We also want to give an insight regarding other factors that may affect mental workload such as driving experience. The specific aims of this experiment as follows:

1. To validate the hypothetical relationship between mental workload and performance with experimental data.
2. To acquire a better understanding of mental workload theory with consideration of internal and external factors.
3. To evaluate the effects of **external factors** of mental workload to driver's mental workload.
4. To assess the effects of **experience in driving** to driver's mental workload.

5.2 Hypotheses

Our hypotheses for this experiment as follows:

1. The combination of internal and external factors will increase driver's mental workload.
2. The decreasing value of SDLP does not mean the improvement in performance.
3. The external factors will give more effects to driver's mental workload
4. The experience of driving will give different results on handling with workload from internal or external factors

5.3 Method

5.3.1 Participants

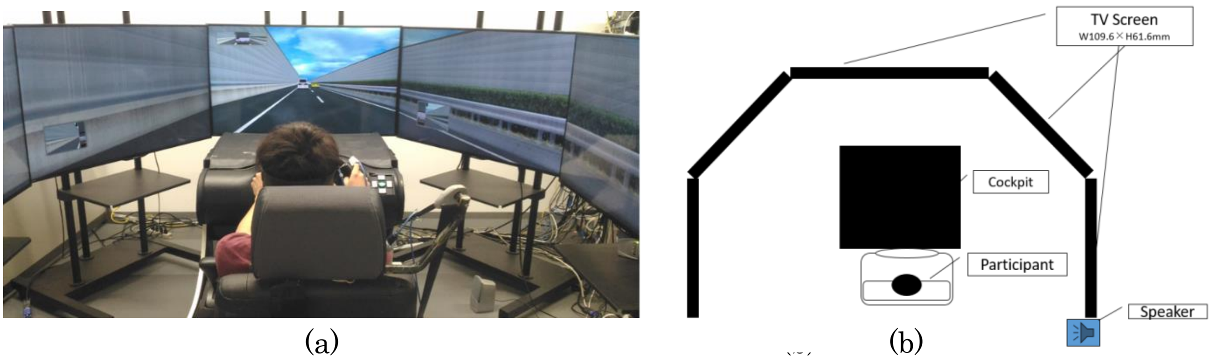
Overall 55 male participants participated in this experiment. During the experiment, five participants withdrew because of simulator sickness. From 50 participants divided into two groups. 1) Novice and 2) Experienced. They recruited based on these criteria:

- Novice: The person who is in age between 18 Years old to 25 years old. Has a valid driving license for less than five years.
- Experienced: The person who is in age above 25 years old and has a valid driving license for more than five years.

For novice group, 28 Participants with mean age: 21.5 Years old, SD: 1.4. Their average driving experience was 2.5 Years, SD: 1.14. While for the Experienced group, total 22 participants were recruited. Their mean age was 31.4 Years old, SD: 4.6 and their average driving experience was 12.1 Years, SD: 4.63.

5.3.2 Apparatus

A fixed-base driving simulator developed by Mitsubishi Precision, Inc was used (Figure 5.1). It consists of 5 TV screens (H109.6xW61.6mm). The location of screen projectors as shown in Figure 5.1 (b). Blood Volume Pulse (BVP) data were recorded with the Nexus-10 device. The software on Mind media was used to analyze data of HR and HRV. BVP device attaches to a finger of participants.



(a) (b)

Figure 5.1 The driving simulator used in the study

5.3.3 Tasks

5.3.3.1 Main Task (External Factor)

The main task was to drive safely in the left lane. As we mentioned before, for this experiment, we want to focus on the effects of external factors to driver mental workload. In all traffic conditions, participants were asked to follow a lead vehicle located in front of the host vehicle, and the following vehicle was located behind the host vehicle as shown in Figure 5.2. If they meet with a crash during a trial, participants would have to start a new trial all over again. This way, participants will try their best for not being involved in a crash.

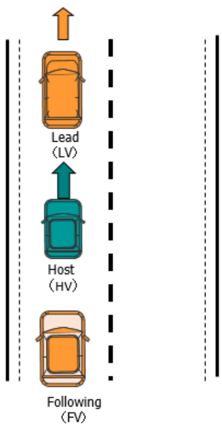


Figure 5.2 Location of LV and FV

Table 5.1 Specification of Traffic Conditions

No.	Traffic Condition		Traffic Geometry	Width of cruising lane (m)	Road Layout	Speed (Km/h)
1	Non-Hazardous Condition	NHC	Straight Pathway	3.5	Highway	85 (constant speed)
2	Narrow Road Driving	NRD	Straight Pathway	2.75	Highway	85 (constant speed)
3	Curve Road Driving	CRD	250m in radius repeatedly curve	3.5	Highway	85 (constant speed)
4	Hazardous Condition	HC	Straight Pathway	3.5	Highway	between 20-65
5	Urban-Hazardous Condition	UHC	Straight Pathway	3.5	Urban	between 20-65

Table 5.1 shows the specification of traffic conditions been used in this experiment. As shown in the table, five traffic condition has been designed in this experiment.

Non-Hazardous Condition (NHC)

The driving course for Non-Hazardous Condition was the same traffic condition as in Experiment#1 and Experiment#2. The road was designed to be a straight pathway with approximately 10km distance with 3.5m width lane. In this traffic condition, participants drive in a highway traffic environment and the both Lead and following vehicle cruise with 85km/h constant speed throughout 7 minutes of a trial.

Narrow Road Driving (NRD)

For Narrow Road Driving (NRD) condition, participants need to drive in a highway traffic condition, and both lead and following vehicle drove with 85km/h. The width of the road for this traffic condition was reduced from 3.5 to 2.75m.

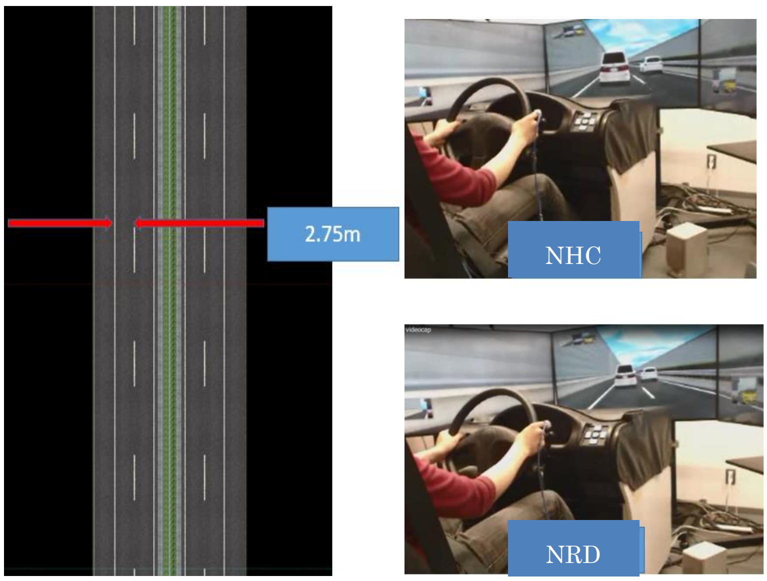


Figure 5.3 Narrow Road Driving (NRD)

Figure 5.3 shows the experiment environment and the different width between NHC and NRD traffic conditions.

Curve Road Driving (CRD)

For curve driving, the cruising speed, the environment and the width of the road were the same as the NHC. The difference was only in the geometry of the road where participants have to experience a continuous curved with a radius of 250m throughout the 7 minutes of a trial.

Figure 5.4 shows the road geometry and environment for the curve road driving condition.

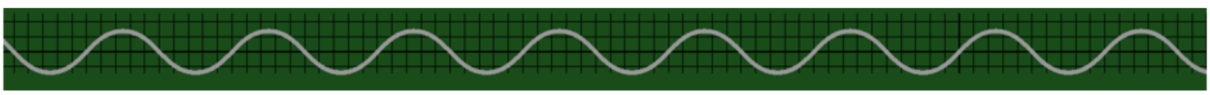


Figure 5.4 Curve Road Driving Geometry and Environment

Hazardous Condition (HC)

Under HC condition, participants were also asked to maintain the distance between LV and FV. At this time, both LV and FV cruised between the speed of 20km/h and 65km/h. Intentionally LV and FV performed an abrupt braking of 0.35g and also a quick acceleration. Time to make a sudden braking and rapid acceleration were determined at random (approximately twice the speed of the changes for each 400-meter run). The participants were requested to ensure a safe following distance and to be careful to an abrupt variation of speed on both FV and LV.

Urban Hazardous Condition (UHC)

In Urban Hazardous Condition, the traffic condition and the movement were the same with HC. However, we change the environment from driving on highway course to drive on the urban road. Figure 5.5 shows the traffic environment and the view from the experiment.

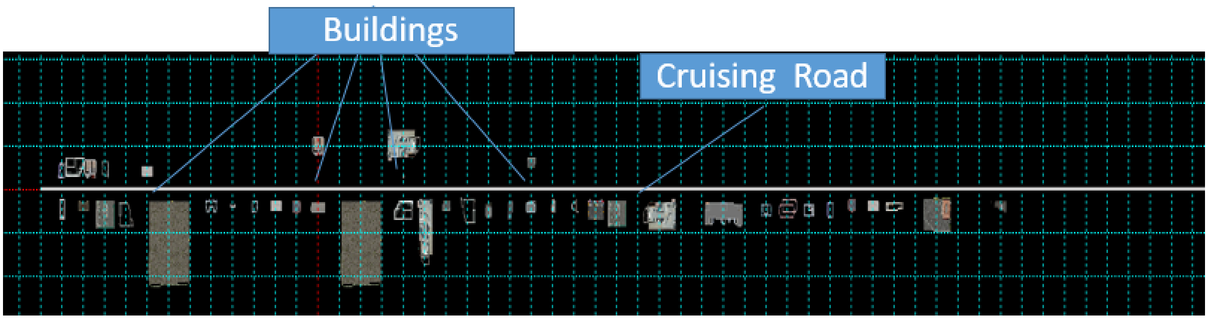


Figure 5.5 Urban Hazardous Condition (UHC)

5.3.3.2 Secondary Task (Internal Factor)

For the internal factor of mental workload, a secondary task has been used. Participants were requested to carry out a two-minute Mathematical Arithmetic Task (MAT) in a 7-minute trial, 150 seconds after the start and finished in 270 seconds of a trial as shown in Figure 5.6. This consideration has been made to decrease the noise due to the start of a trial. The data of the 30s at the start and 30s before the end was regarded as a buffer time. For this experiment, we only used MAT2 (mathematical arithmetic task with two-digit numbers- the difficult level of MA task). The same with has been used in the previous experiment. NMAT indicates as no mathematical arithmetic task being performed in this experiment.

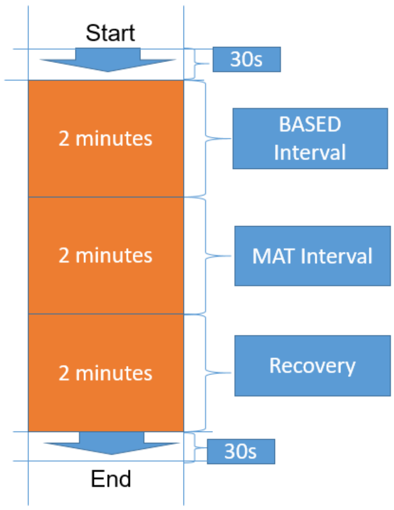


Figure 5.6 Time period for one trial

5.3.4 Experimental Design

The tasks were designed using two manipulations in a within-subject factors factorial design: Secondary Task (NMAT, or MAT2) and Traffic Condition (NHC, Width, Curve, HC, or Urban Hazardous).

5.3.5 Procedure

Upon arriving at the venue of the experiment, participants were explained regarding the experiment and signed a consent letter. They were given time to get familiarized with the simulator and also the MAT2. Before starting the experiment, BVP sensor device was attached to the right-hand finger of participants, while ECG to be attached to the chest of participants, and the Headband BVP sensor was attached to the forehead of participants. As shown in Table 5.2, every participant performed ten types of combinations between secondary task and the traffic condition. As a subjective measurement, participants were requested to answer NASA-TLX [27] and RSME [64] immediately after finishing every trial. We changed the time from 6 minutes in a trial to 7 minutes as shown in Figure 5.6.

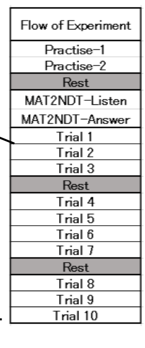
In this experiment, MAT2NDT was the condition when the participants only performed MAT2 under Non-Driving Task. They had to sit on the steering wheel and were given numbers to solve MAT2, first two minutes they were requested to solve the calculation by not uttering it. They just need to hear the number being produced by the speaker. After finishing the first two minutes, then they were asked to answer the MAT2 verbally.

Our intention was to find if there was an effect on physiological measurements of MAT2 by just thinking in their mind without verbally answer it with they do both the calculation and verbally answer it.

Table 5.2 Experiment Procedure

Randomly Arranged

No.	Secondary Task	Traffic Condition	Driving Condition
1	NMAT	NHC	NMATNHC
2	MAT2	NHC	MAT2NHC
3	NMAT	NRD	NMATNRD
4	MAT2	NRD	MAT2NRD
5	NMAT	CRD	NMATCRD
6	MAT2	CRD	MAT2CRD
7	NMAT	HC	NMATHC
8	MAT2	HC	MAT2HC
9	NMAT	UHC	NMATUHC
10	MAT2	UHC	MAT2UHC



<u>Training Session:</u>

For the training session, we designed a course that including four types of road conditions. Each of the road conditions took approximately 2 minutes to be tested, and they were connected each other. In the beginning, they were asked to follow the lead vehicle and drove in a Baseline Driving road condition, in a highway environment and

3.5m width lane. After finishing arriving at the end of the road, they drove in the same highway environment, but with a narrower path of 2.75 m. After reaching the end of the road, they tried to drive in curve road condition in the highway environment and finally, they drove in an urban road environment.

During this training session, they would hear two types of warning message intended to help them performing better in the real trial.

1. "中央車線を目安に走行してください" translated as "Please drive in the middle of the lane as a reference" if they were stepping out from the lane.

2. "速度をあげてください" translated as "Please increase your speed" if the distance between Lead Vehicle and Host Vehicle is larger than 30 meters.

5.3.6 Dependent Variables

Below is the list of Dependent Variables tested in this experiment while the independent variables were the Driving conditions and the age of participants?

Measurement Groups	Dependent Variables
Physiological	HR and HRV From Finger BVP
	HR and HRV From Head BVP
	HR and HRV From ECG
Performance	Standard Deviation of Lateral Position
	Steering Entropy
Subjective	NASA-TLX
	RSME

5.3.7 Data Analysis

To increase the accuracy of estimation the levels of mental workload as we discussed in chapter 3, we normalized physiological and performance data. This method is useful to find a standard value that indicates the same level of mental workload. It was also to ensure that data were not affected by "state-related effort" especially fatigue, e.g. considering the data maybe has already increased from the beginning of one trial and not only during performing MAT.

Normalization equation as shown in equation (2). In a 7-minute trial, minute 2.30 to minute 4.30 from the beginning was regarded as MAT Period data.

We set two minutes of 30 seconds after starting a trial as a 'Base Interval' and compared the data to MAT period from 150s to 270s of a trial. Results from Experiment#1 were also used to determine the base interval which is 120s (the same length with MAT period data). The interval between 30 seconds to 150 seconds was selected as base interval data.

$$Normalization = \left(\frac{MAT\ Period\ Data}{Base\ Interval\ Data} \times 100\right)\% \qquad (2)$$

General Linear Model for Repeated-measures analysis in SPSS (version 21.0) was used to analyze the data. Three independent variables: Age of participants (two groups), types of secondary Task (NMAT and MAT2) and types of traffic conditions (NHC, NRD, CRD, HC, and UHC) were analyzed by using a 2 X 2 X 3 mixed factorial repeated-measures design.

5.4 Results

5.4.1 Non-Driving Task

A repeated-measures ANOVA was conducted with Sources (Finger BVP, Head BVP, and ECG) and Group (Novice and Experienced) as between-subjects factors. A 2x2x2 mixed factorial designed was used with HR during 2 minutes of "Listening" to the MAT2 and 2 Minutes of actually "answer" the task verbally as within-subjects factors. Both of these tasks were performed under non-Driving Task.

As shown in Figure 5.7, there were significant differences between Listen and Answer, $F(1,144)=0.258$, $p=0.612$, $\eta_p^2=0.002$. The differences between Groups and Sources were not significant, $F(1,144)=0.258$, $p=0.612$, $\eta_p^2=0.002$ and $F(2,144)=0.027$, $p=0.973$, $\eta_p^2=0.000$ respectively. The interaction between those factors also not significant.

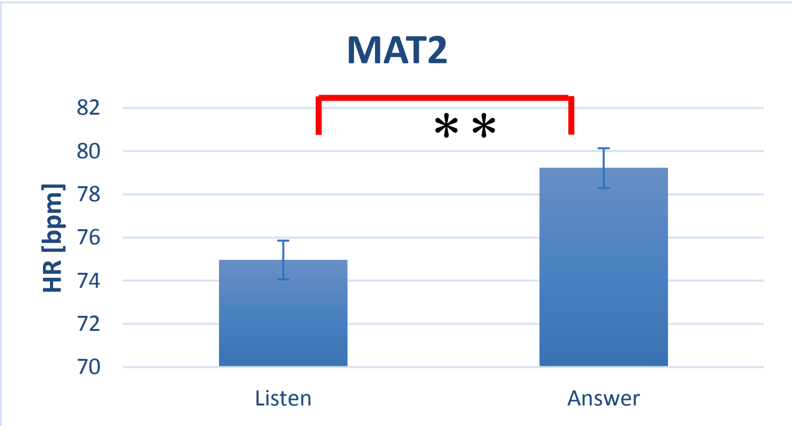


Figure 5.7 Results of MAT2NDT

These results show that the secondary task they performed was different between merely listening to the task without answering it with actually answering it verbally. It gives us an insight on the designing of a secondary task. It also gives us information about the physiological data that we acquired may be affected by the verbal response to the task. However, as we cannot confirm whether while listening, they were actually 'think' to answer to the MAT2 or not. We make a consideration about the results of physiological measurements presented in this experiment were mixed between response as a result of "<u>think</u> to answer the task" and "answer the task <u>verbally</u>".

5.4.2 Physiological Measurement

5.4.2.1 HR and Normalized HR

An analysis of variance with repeated measures was computed from the individual means of HR and Normalized HR in five traffic conditions (Baseline, Narrow, Curve, Hazardous and Urban Hazardous). A 3x2x2x5 mixed factorial design was used. Three type of sources (Finger BVP, ECG or Head BVP) and two driving experience groups (Novice or Experienced) as between-subjects factors.Two types of secondary task (NMAT or MAT2) and five traffic conditions (Baseline, Narrow, Curve, Hazardous Condition and Urban Hazardous Condition) as within-subjects factors.

Table 5.3 Univariate test results for HR and Normalized HR. MAT= Type of a secondary task, NMAT and MAT2, Sources= Sources of data, Finger BVP, Headband BVP and ECG, Group: Novice and Experienced, TC: Traffic Conditions, NHC, Narrow, Curve, Hazardous, and Curve Hazardous. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when Mauchly's test showed a violation of sphericity.

Effect	HR				Normalized HR			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**1, 144**	**133.994**	**<0.001**	**0.482**	**1, 144**	**133.994**	**<0.001**	**0.513**
Sources (S)	2,144	0.26	0.772	0.004	2,144	0.420	0.658	0.006
Group (G)	1,144	0.062	0.803	0.000	1,144	0.053	0.818	0.000
MAT x S	2,144	0.081	0.922	0.001	2,144	0.081	0.926	0.001
MAT x G	1,144	0.074	0.786	0.001	1,144	0.074	0.707	0.001
MAT x S x G	2,144	0.152	0.859	0.002	2,144	0.152	0.889	0.002
Traffic Condition (TC)	2.6, 373.6	0.884	0.437	0.024	3.1, 439.8	1.126	0.339	0.008
TC x S	5.2, 373.6	0.344	0.892	0.005	6.1, 439.8	0.303	0.937	0.004
TC x G	2.6, 373.6	2.21	0.096	0.015	3.1, 439.8	1.527	0.206	0.010
TC x S x G	5.2, 373.6	0.275	0.931	0.004	6.1, 439.8	0.150	0.990	0.002
MAT x TC	3.3, 471.0	2.029	0.103	0.014	**3.2, 466.2**	**2.872**	**0.032**	**0.020**
MAT x TC x S	6.5, 471.0	0.305	0.944	0.004	6.5, 466.2	0.268	0.959	0.004
MAT x TC x G	3.3, 471.0	0.61	0.623	0.004	3.2, 466.2	0.537	0.671	0.004
MAT x TC x S x G	6.5, 471.0	0.087	0.998	0.001	6.5, 466.2	0.084	0.999	0.001

Table 5.3 shows univariate test results for HR and Normalized HR data. According to the results of repeated-measures ANOVA, the significant main effect of the type of secondary tasks was on both HR and Normalized HR, $F(1,144)=133.994$, $p<0.001$, $\eta_p^2=0.482$ and $F(1,144)=133.994$, $p<0.001$, $\eta_p^2=0.513$ respectively. The interaction between MAT and Traffic conditions was also found statistically significant for the Normalized HR data, $F(3.2,466.2)=2.872$, $p<0.032$, $\eta_p^2=0.020$.

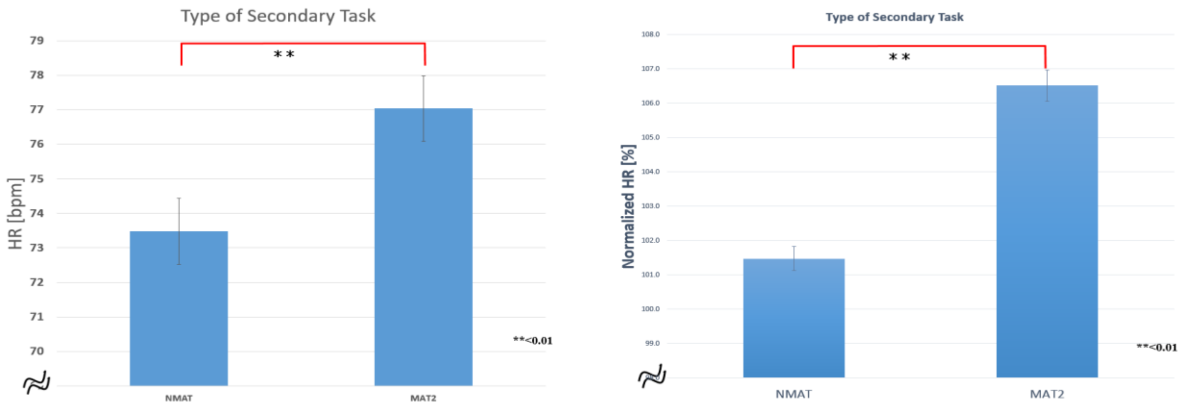


Figure 5.8 Mean HR based on type of secondary task

Figure 5.8 shows graphs of the average value of HR and Normalized HR according to the type of secondary task. As shown in the figure, HR during MAT2 was higher than during no mathematical arithmetic task. This result is consistent with our previous experiment which HR and Normalized HR were higher when accomplishing the MAT2 type of secondary task.

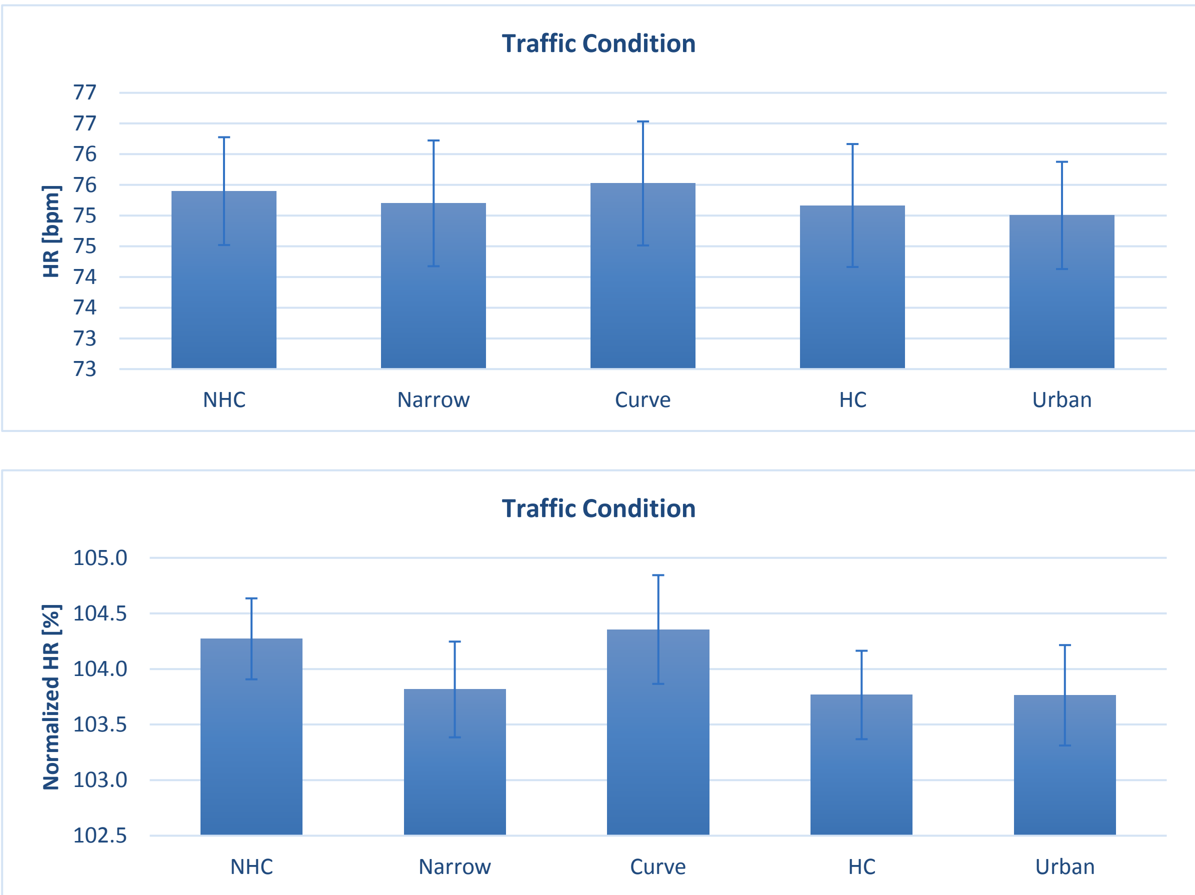


Figure 5.9 Mean HR and Normalized HR according to the Traffic conditions

Figure 5.9 shows graphs of the mean value for HR and Normalized HR data according to the type of Traffic Conditions. Analysis of repeated ANOVA's results indicate that the differences between the traffic conditions was not significant, HR: $F(2.6, 373.6)=0.884$, $p=0.437$, $\eta_p^2=0.024$, Normalized HR : $F(3.1, 439.8)=1.126$, $p=0.339$, $\eta_p^2=0.008$.

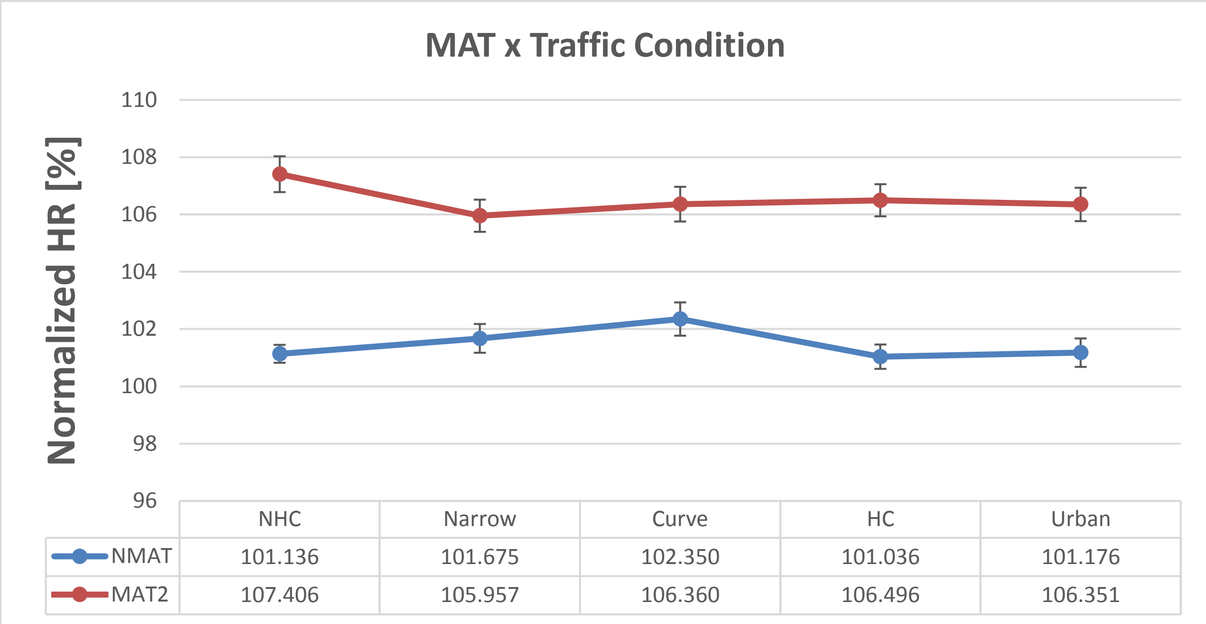


Figure 5.10 The interaction between MAT and Traffic Condition

For Normalized HR data, the interaction between MAT and Traffic conditions was found significant, $F(3.2,466.2)=2.872$, $p=0.032$, $\eta_p^2=0.020$. The graph in Figure 5.10 illustrates the interaction. The value of Normalized HR on all traffic conditions was higher when participants were accomplishing the secondary task compared to NMAT. It also can be seen here that during NMAT, HR was higher in curve road traffic condition compared to baseline driving. However, when accomplishing the MAT2 the trend did not exist anymore with NHC driving gives the highest value among others (107.41%).

5.4.2.2 HRV and Normalized HRV

The same analysis as HR and Normalized HR was performed to HRV and Normalized HRV data.

Table 5.4 shows univariate test results for HRV and Normalized HRV data. According to the results of repeated-measures ANOVA, the significant main effect of the type of secondary tasks was found on HRV data, $F(1,144)=8.287$, $p=0.005$, $\eta_p^2=0.054$. Significant main effects of sources of data also found for HRV data, $F(2,144)=7.875$, $p=0.001$, $\eta_p^2=0.099$. Another significant main effect was found for Normalized HRV on the interaction between Traffic conditions and Groups of participants, Novice and Experienced, $F(2.2,309.5)=4.060$, $p<0.016$, $\eta_p^2=0.027$.

Table 5.4 Univariate test results for HRV and Normalized HRV. MAT= Type of a secondary task, NMAT and MAT2, Sources= Sources of data, Finger BVP, Headband BVP and ECG, Group: Novice and Experienced, TC: Traffic Conditions, NHC, Narrow, Curve, Hazardous, and Curve Hazardous. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when Mauchly's test showed a violation of sphericity.

Effect	HRV				Normalized HRV			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**1, 144**	**8.287**	**0.005**	**0.054**	1, 144	0.397	0.530	0.003
Sources (S)	**2,144**	**7.875**	**0.001**	**0.099**	2,144	0.208	0.813	0.003
Group (G)	1,144	0.000	0.989	0.000	1,144	0.405	0.526	0.003
MAT x S	2,144	0.139	0.870	0.002	2,144	0.802	0.450	0.011
MAT x G	1,144	1.678	0.197	0.012	1,144	0.633	0.428	0.004
MAT x S x G	2,144	0.108	0.898	0.001	2,144	0.153	0.858	0.004
Traffic Condition (TC)	3.3, 470.1	1.887	0.126	0.013	2.2,309.5	1.763	0.170	0.012
TC x S	6.5, 470.1	0.602	0.743	0.008	4.3, 309.5	1.611	0.167	0.022
TC x G	3.3, 470.1	0.984	0.405	0.007	**2.2,309.5**	**4.060**	**0.016**	**0.027**
TC x S x G	6.5, 470.1	0.264	0.961	0.004	4.3, 309.5	1.079	0.369	0.015
MAT x TC	4,141	0.397	0.810	0.011	2.5,360.4	1.532	0.212	0.011
MAT x TC x S	8,284	0.560	0.810	0.016	5.0,360.4	0.592	0.707	0.008
MAT x TC x G	4,141	0.734	0.570	0.020	2.5,360.4	1.235	0.296	0.009
MAT x TC x S x G	8,284	0.197	0.991	0.006	5.0,360.4	0.385	0.860	0.005

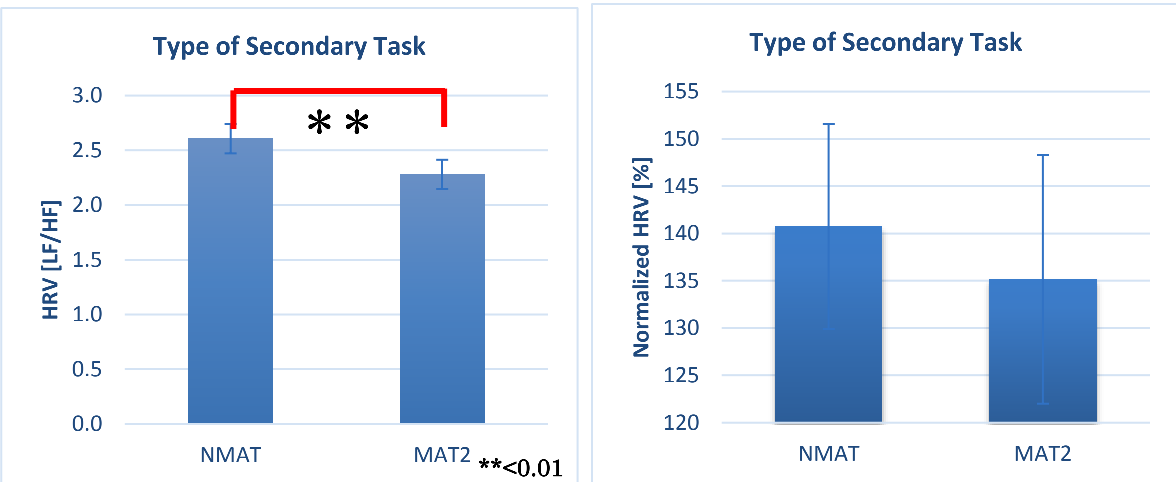
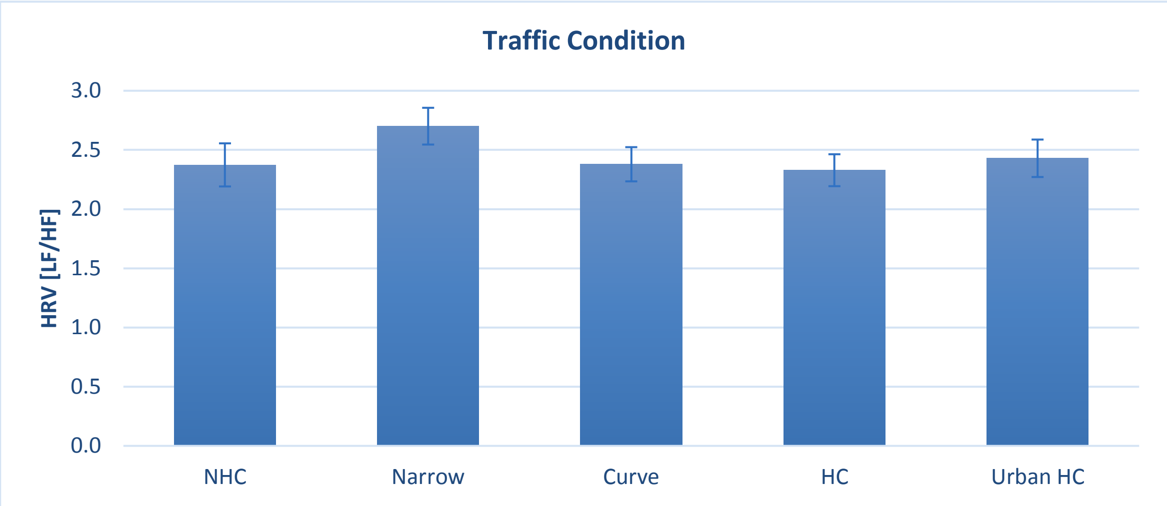


Figure 5.11 Mean value of HRV according to the type secondary task

Figure 5.11 shows the mean value graphs of HRV and Normalized HRV according to the type of secondary task. From the graphs, HRV was reduced significantly (p<0.001) when participants were accomplishing the MAT2. This reduction was not found during our previous experiment. The existence of secondary task was reduced the HRV value.



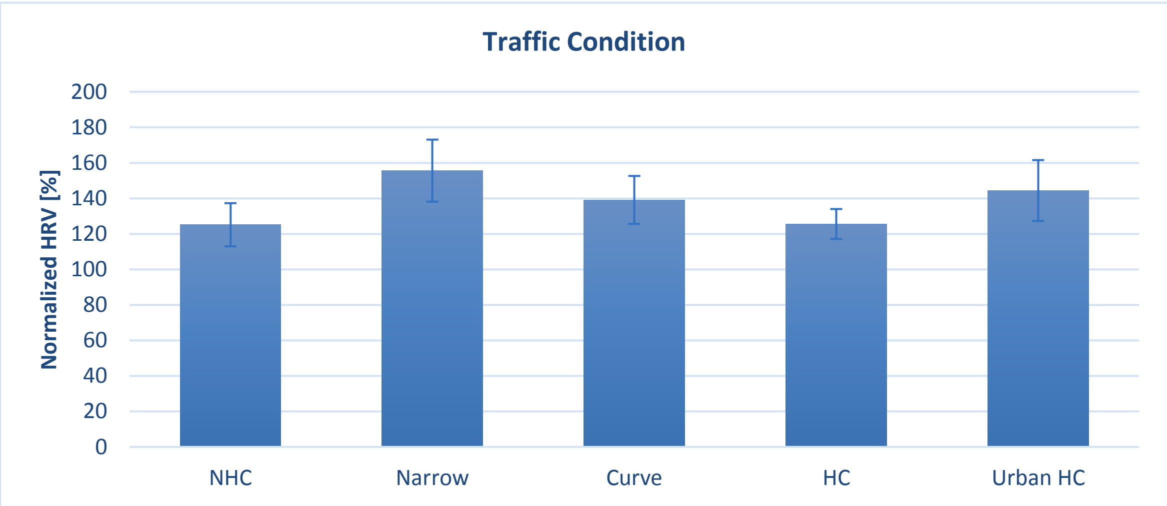


Figure 5.12 Mean value of HRV according to Traffic conditions

Figure 5.12 shows the average value of HRV and Normalized HRV according to traffic conditions. From the graph, the lowest value of HR was found on Hazardous Condition (2.33), while NHC recorded the lowest for Normalized HRV, 125.1%. The results of repeated ANOVA, however, reveal that the differences were not significant.

5.4.3 Performance Measurements

5.4.3.1 Primary task

a) Standard Deviation of Lateral Position (SDLP)

An analysis of variance with repeated-measures was computed from the individual means of SDLP and Normalized SDLP in five traffic conditions (Baseline, Narrow, Curve, Hazardous and Urban Hazardous). A 2x2x5 mixed factorial design was used, driving experience group (Novice or Experienced) as between-subjects factors and type of secondary task (NMAT or MAT2) and traffic conditions (NHC, Narrow, Curve, Hazardous Condition and Urban Hazardous Condition) as within-subjects factors.

Table 5.5 Univariate test results for SDLP and Normalized SDLP. MAT= Type of a secondary task, NMAT, and MAT2, Group: Novice and Experienced, TC: Traffic Conditions, NHC, Narrow, Curve, Hazardous, and Curve Hazardous. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when Mauchly's test showed violation of sphericity

Effect	SDLP				Normalized SDLP			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**1,48**	**101.118**	**<0.001**	**0.678**	**1,48**	**117.830**	**<0.001**	**0.711**
Group (G)	**1,48**	**7.851**	**0.007**	**0.141**	1,48	0.343	0.561	0.007
MAT x G	1,48	3.469	0.069	0.067	1,48	0.236	0.629	0.005
Traffic Condition (TC)	**1.9, 93.1**	**159.098**	**<0.001**	**0.768**	**1.7, 80.4**	**135.308**	**<0.001**	**0.738**
TC x G	**1.9,93.1**	**5.505**	**0.006**	**0.103**	1.7, 80.4	2.024	0.146	0.040
MAT x TC	**3.2, 155.4**	**5.880**	**0.001**	**0.109**	**3.3, 158.3**	**4.424**	**0.004**	**0.084**
MAT x TC x G	3.2, 155.4	0.461	0.725	0.010	3.3, 158.3	0.231	0.891	0.005

Table 5.5 shows univariate test results for SDLP and Normalized SDLP data. According to the results of repeated-measures ANOVA, significant main effect of the type of secondary tasks was found on both SDLP and Normalized SDLP data, $F(1,48)=101.118$, $p<0.001$, $\eta_p^2=0.678$ and $F(1,48)=117.830$, $p<0.001$, $\eta_p^2=0.711$, respectively. The effects of Traffic conditions were also significant for both SDLP and Normalized SDLP, $F(1.9,93.1)=159.098$, $p<0.001$, $\eta_p^2=0.768$ and $F(1.7, 80.4)=135.308$, $p<0.001$, $\eta_p^2=0.738$, respectively. The interaction between MAT and Traffic Conditions was also found significant on SDLP and Normalized SDLP, SDLP: $F(3.2,155.4)=5.880$, $p=0.001$, $\eta_p^2=0.109$ and Normalized SDLP: $F(3.3,158.3)=4.424$, $p=0.004$, $\eta_p^2=0.084$. Significant main effects of groups was also found for SDLP data, $F(1,48)=7.851$, $p=0.007$, $\eta_p^2=0.141$. Another significant main effect was found for Normalized SDLP on the interaction between Traffic conditions and Groups of participants, Novice and Experienced, $F(3.2,155.4)=5.880$, $p=0.001$, $\eta_p^2=0.109$.

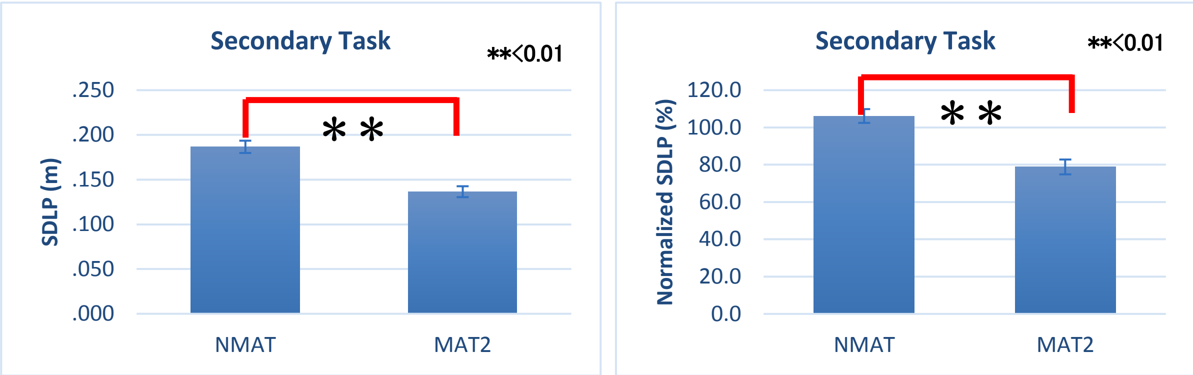


Figure 5.13 Mean value for SDLP and Normalized SDLP according to the type of secondary task

Figure 5.13 shows a mean value of SDLP and Normalized SDLP for the type of secondary task. According to graphs, both SDLP and Normalized SDLP was significantly reduced when participants were accomplishing secondary tasks. This phenomena could be interpreted as participants less swerving when accomplishing a secondary task. This finding is consistent with our previous findings from Experiment#1 and Experiment#2.

Figure 5.14 Mean value of SDLP among Traffic Conditions

Figure 5.14 shows the mean value of SDLP among the traffic conditions. As shown in the figure, SDLP was the highest under Curve conditions, and the differences were significant among other traffic conditions. The lowest value for SDLP was recorded from Hazardous traffic conditions with 0.122m. The difference between Baseline and Hazardous Conditions was also found significant ($p<0.001$).

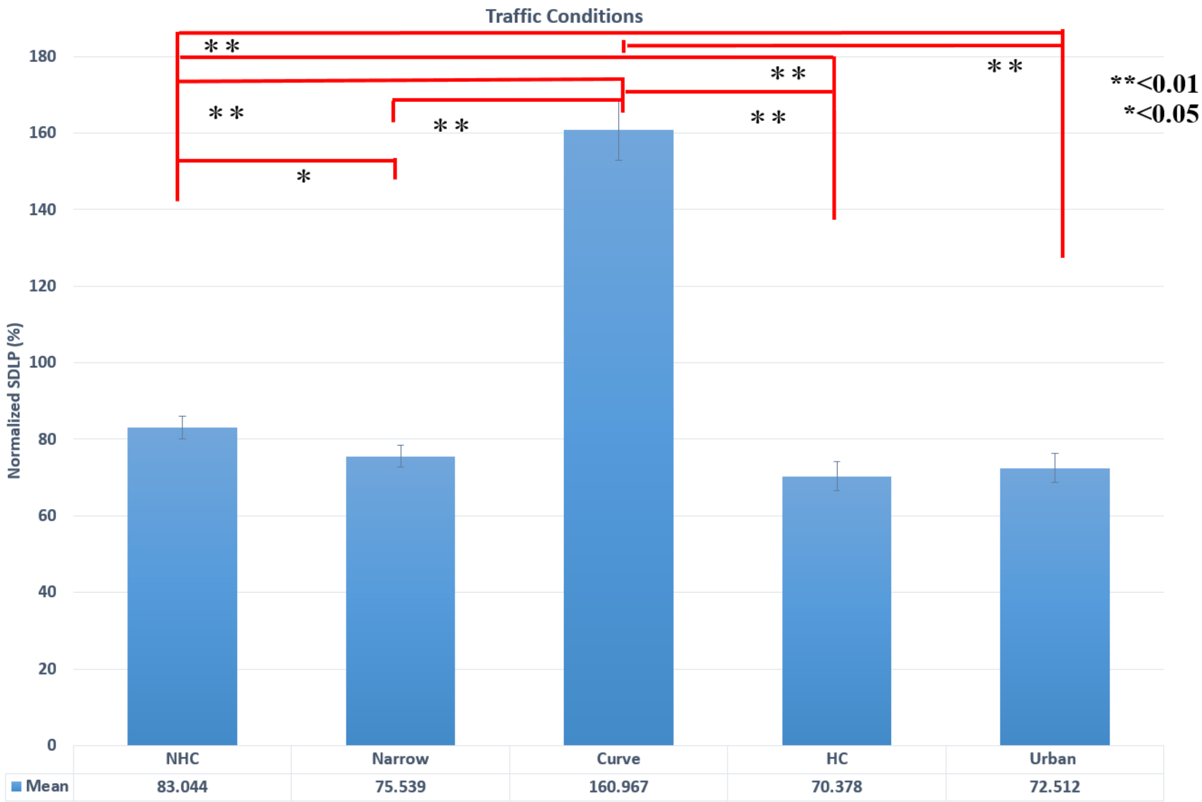
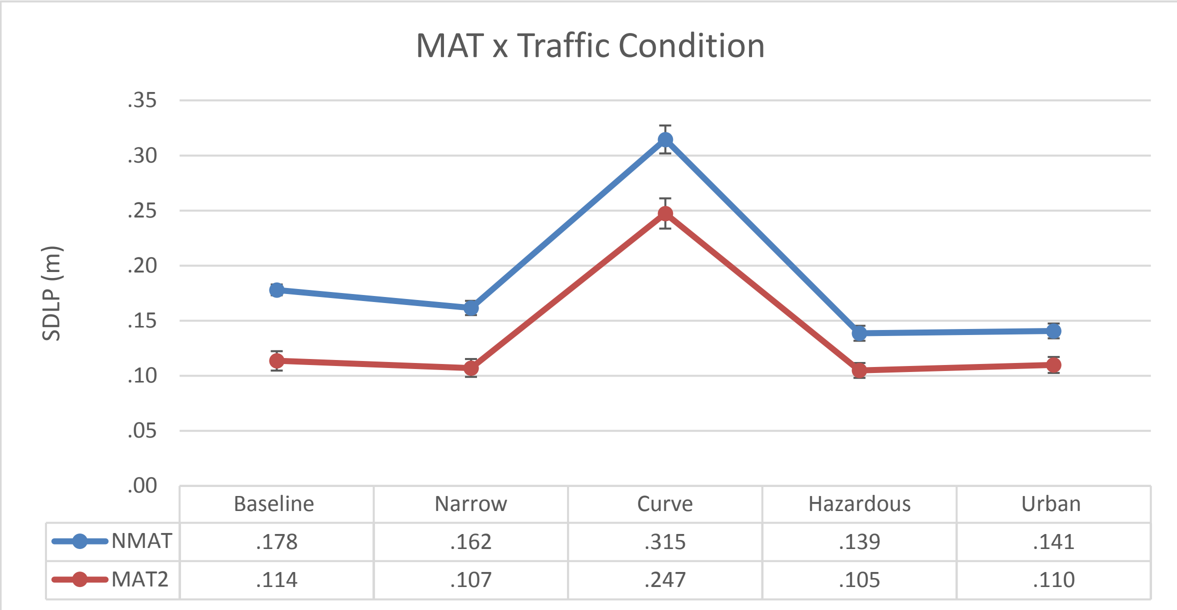
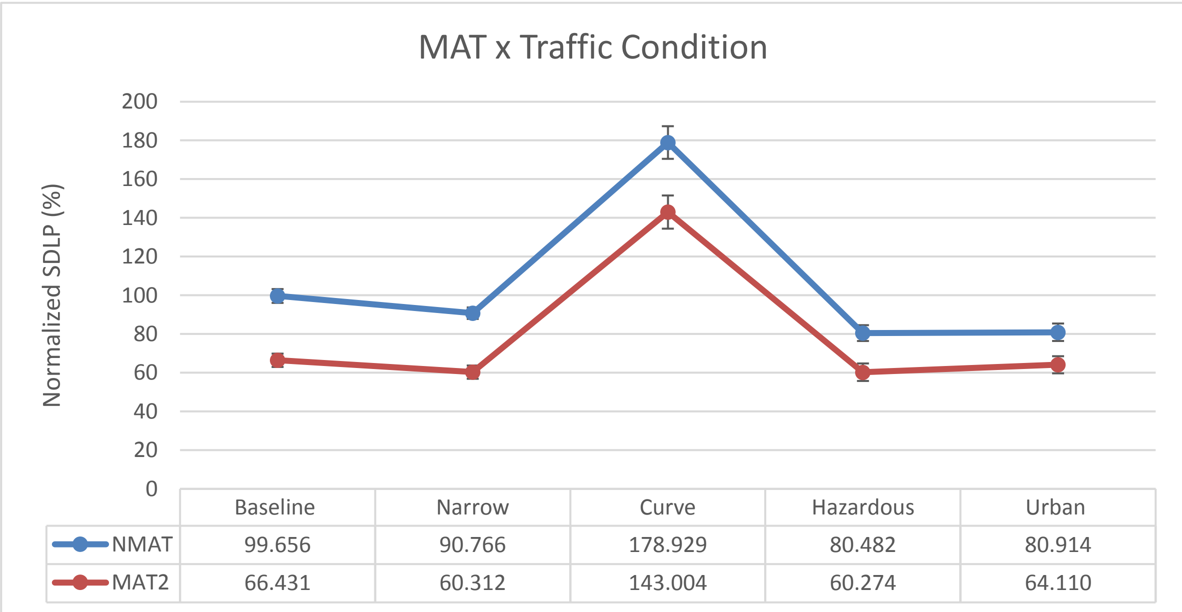


Figure 5.15 Mean Normalized SDLP among traffic conditions

Figure 5.15 shows the average value of Normalized SDLP among the traffic conditions. As shown in the graph, Normalized SDLP was still the highest under Curve conditions, and the differences were significant among other traffic conditions. The lowest value for Normalized SDLP was recorded from Hazardous traffic conditions. The difference between Baseline and Hazardous Conditions was also found significant ($p<0.001$).



(a)



(b)

Figure 5.16 Interaction between MAT and Traffic Conditions

Figure 5.16 (a) shows a graph of interaction between MAT and Traffic Conditions. As shown in the graph the mean of SDLP was higher in all traffic conditions when participants were accomplishing the secondary task of MAT2. Figure 5.16 (b) also shows a similar trend for Normalized SDLP data.

b) <u>Steering entropy</u>

As for steering entropy data, a similar analysis as SDLP has been performed on the data.

Table 5.6 Univariate test results for Steering Entropy and Normalized Steering Entropy. MAT= Type of a secondary task, NMAT, and MAT2, Group: Novice and Experienced, TC: Traffic Conditions, NHC, Narrow, Curve, Hazardous, and Curve Hazardous. Significant effects (p<0.05) are shown in bold.

Effect	Steering Entropy				Normalized Steering Entropy			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**1,48**	**154.852**	**<0.001**	**0.763**	**1,48**	**149.918**	**<0.001**	**0.757**
Group (G)	**1,48**	**6.470**	**0.014**	**0.119**	1,48	0.343	0.561	0.007
MAT x G	1,48	0.005	0.946	0.000	1,48	0.007	0.935	0.000
Traffic Condition (TC)	**4,45**	**72.313**	**<0.001**	**0.865**	**4,45**	**70.120**	**<0.001**	**0.862**
TC x G	4,45	2.353	0.068	0.173	4,45	2.384	0.065	0.175
MAT x TC	**4,45**	**4.979**	**0.002**	**0.307**	**4,45**	**5.110**	**0.002**	**0.312**
MAT x TC x G	4,45	0.072	0.990	0.006	4,45	0.081	0.988	0.007

Table 5.6 shows univariate test results for Steering Entropy and Normalized SE data. According to the results of repeated-measures ANOVA, significant main effects of the type of secondary tasks were found on both Steering Entropy and Normalized SE data, $F(1,48)=154.852$, $p<0.001$, $\eta_p^2=0.763$ and $F(1,48)=149.918$, $p<0.001$, $\eta_p^2=0.757$, respectively. The effects of Traffic conditions were also significant for both Steering Entropy and Normalized SE, $F(4,45)=72.313$, $p<0.001$, $\eta_p^2=0.865$ and $F(4,45)=70.12$, $p<0.001$, $\eta_p^2=0.862$, respectively. The interaction between MAT and Traffic Conditions was also found significant on Steering Entropy and Normalized SE, Steering Entropy: $F(4,45)=4.979$, $p=0.002$, $\eta_p^2=0.307$ and Normalized SE:$F(4,45)=5.110$, $p=0.002$, $\eta_p^2=0.312$. Significant main effects of groups were also found for Steering Entropy data, $F(1,48)=6.470$, $p=0.014$, $\eta_p^2=0.119$.

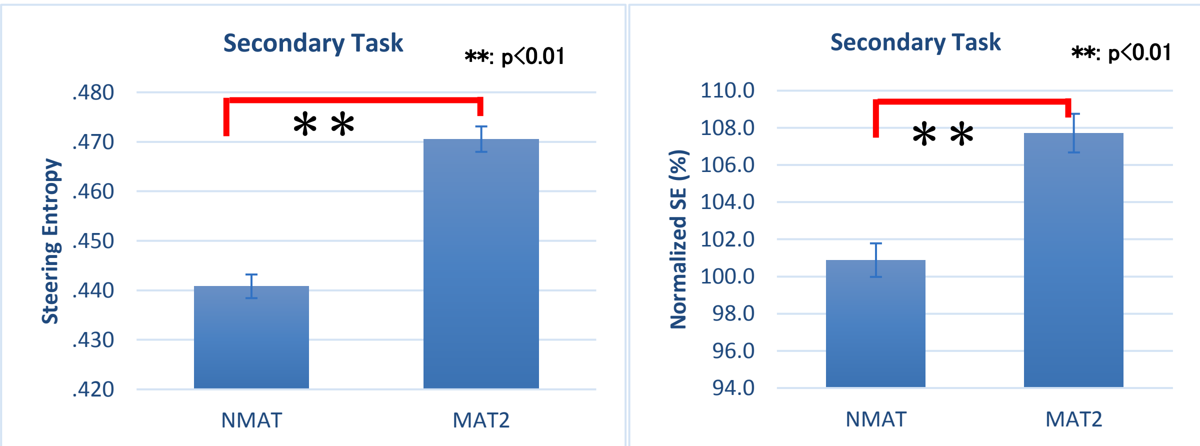


Figure 5.17 Mean value of Steering Entropy and Normalized SE according to traffic conditions

Figure 5.17 shows a mean value of Steering Entropy and Normalized SE according to the type of secondary tasks. According to the graphs, values of Steering Entropy were significantly increased when participants were accomplishing the secondary tasks of MAT2. These findings are also consistent with our finding in Experiment#2.

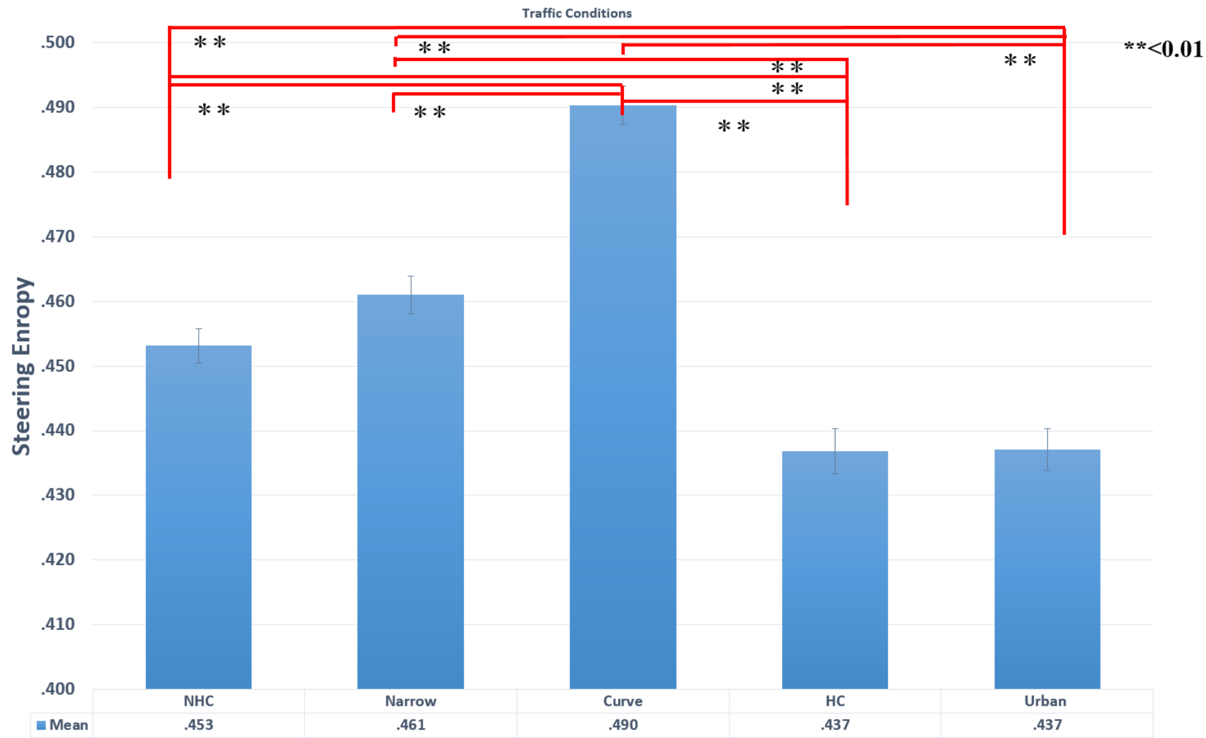


Figure 5.18 Mean value of Steering Entropy among traffic conditions

Figure 5.18 shows the average value of mean steering entropy according to traffic conditions. The mean value was the highest when participants were driving under curve road conditions. The lowest value can be found in the hazardous condition. The mean value of steering entropy also higher when they have to drive in narrow road condition.

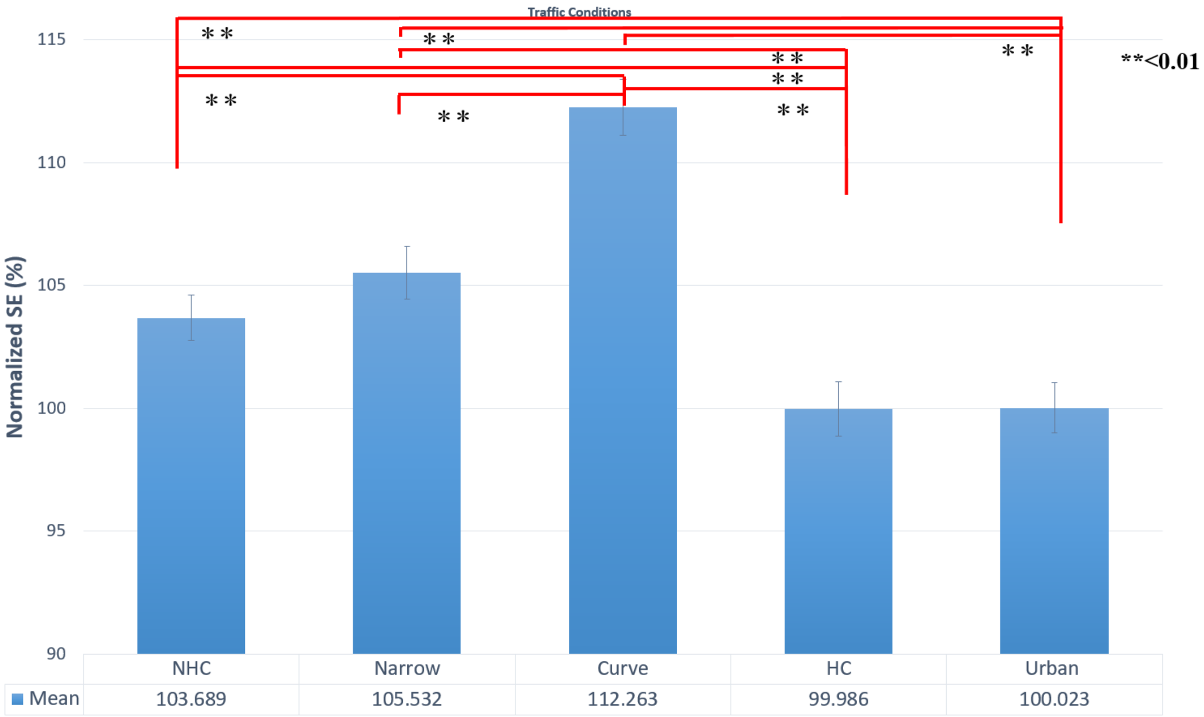
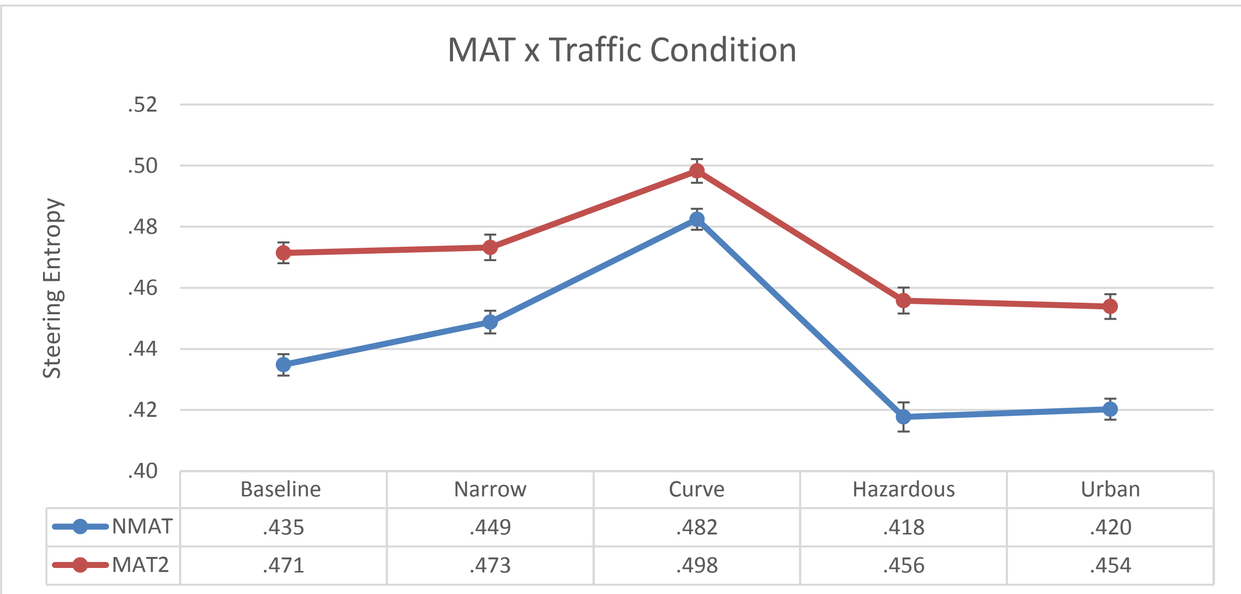
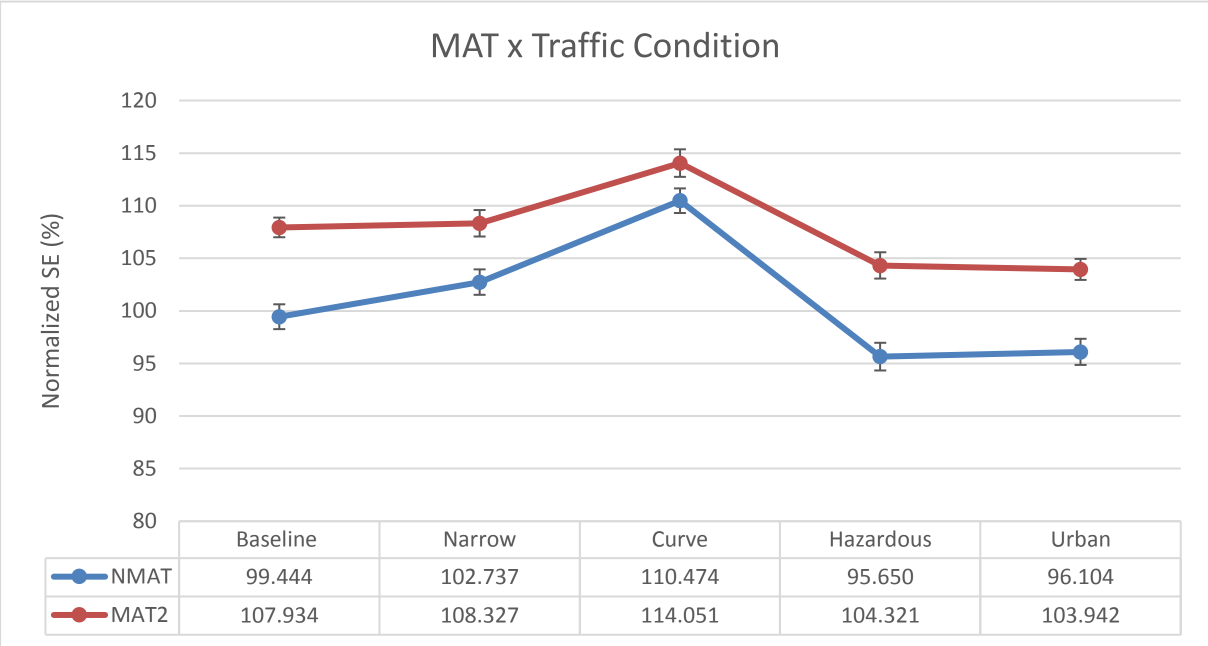


Figure 5.19 Mean value of Normalized SE according to traffic conditions

Figure 5.19 shows an average value of Normalized SE according to traffic conditions. Same pattern with Steering Entropy data, driving during curves traffic conditions gives the highest value of steering entropy. There was also significant differences between curves and other traffic conditions. The lowest value of Normalized SE was found on Hazardous condition (99.99%), almost the same with baseline data.



(a)



(b)

Figure 5.20 Interaction between MAT and Traffic Condition

Figure 5.20 (a) shows a graph of interaction between MAT and Traffic Conditions for Steering Entropy data. As shown in the graph the mean of steering entropy was higher in all traffic conditions when participants were accomplishing the secondary task of MAT2. Figure 5.20 (b) also shows a similar trend for Normalized SE data.

5.4.3.2 Secondary Task

Table 5.7 Univariate test results for Percentage of the Correct answer. Group: Novice and Experienced, TC: Traffic Conditions, NHC, Narrow, Curve, Hazardous, and Curve Hazardous. Significant effects (p<0.05) are shown in bold.

Effect	Percentage of Correct Answer			
	df 1,2	F	p	η_p^2
Traffic Condition (TC)	5,44	1.692	0.157	0.161
Group (G)	**1,48**	**11.214**	**0.002**	**0.189**
MAT x G	5,44	1.211	0.320	0.121

Table 5.7 shows a univariate test results for the percentage of the correct answer. As shown in the table, a significant main effect only found on Groups, between Novice and Experienced.

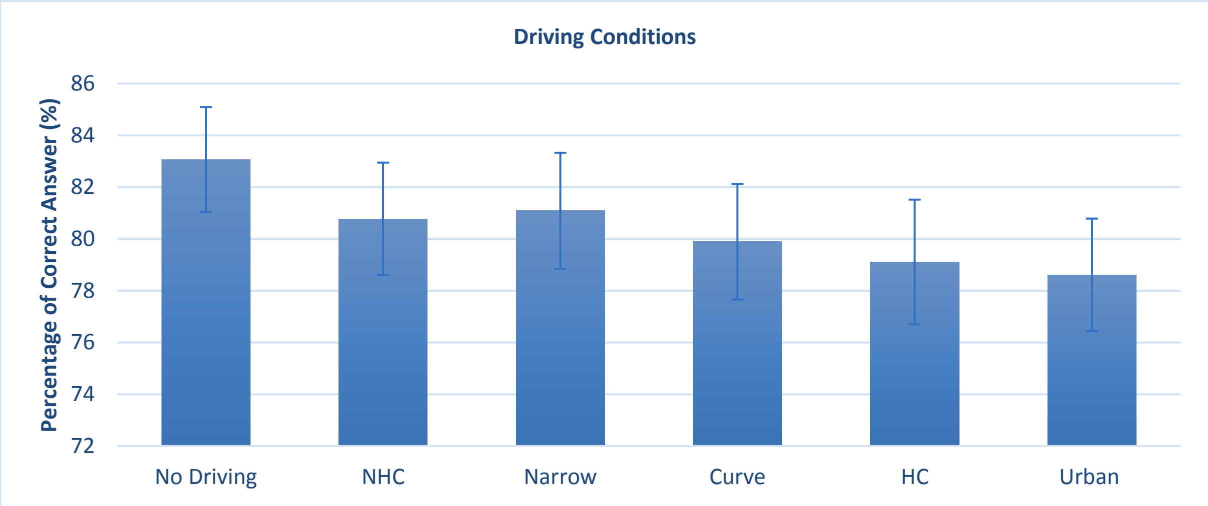


Figure 5.21 Percentage of Correct Answer for Traffic Conditions

Figure 5.21 shows a graph of percentage correct answer for secondary task. The lowest value was when participants were accomplishing the task while under urban HC conditions (78.61%). According to the analysis, the main effects of driving experience was significant on Secondary Task performance, $F(1,48)=11.214$, $p=0.002$, $\eta_p^2=0.189$. As shown in Figure 5.22, Novice Group (87.247) performed better than experienced group (73.547). There was no significant difference found on Traffic conditions $F(5,44)=1.692$, $p=0.157$, $\eta_p^2=0.161$, and the interaction between secondary task and traffic conditions was found insignificant $F(5,44)=1.211$ $p=0.320$, $\eta_p^2=0.121$.

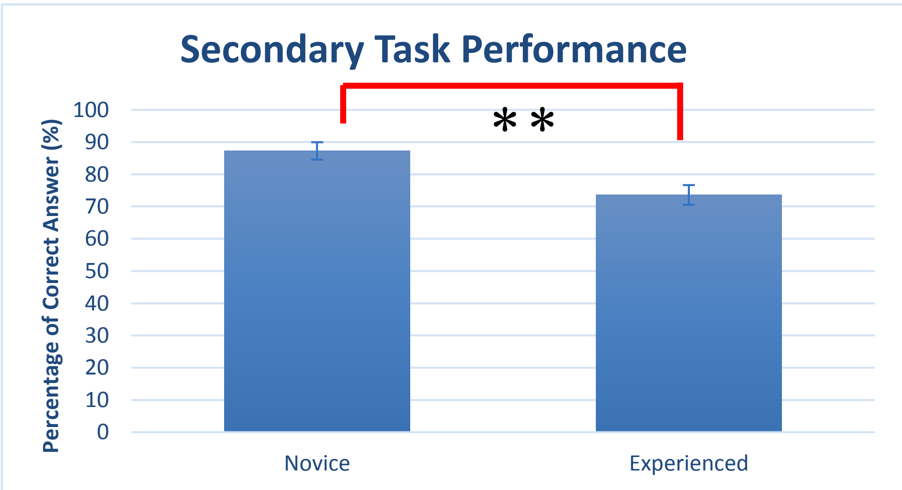


Figure 5.22 Secondary Task performance between groups

5.4.4 Subjective measurements

For subjective measurements, an analysis of variance with a repeated-measures ANOVA was computed from the individual means of NASA-TLX and RSME in five traffic conditions (NHC, Narrow, Curve, Hazardous and Urban Hazardous). A 2x2x5 mixed factorial design was used, with driving experience group (Novice or Experienced) as between-subjects factors and type of secondary task (NMAT or MAT2) and traffic conditions (NHC, Narrow, Curve, Hazardous Condition and Urban Hazardous Condition) as within-subjects factors.

Table 5.8 Univariate test results for NASA-TLX and RSME. MAT= Type of a secondary task, NMAT, and MAT2, Group: Novice and Experienced, TC: Traffic Conditions, NHC, Narrow, Curve, Hazardous, and Curve Hazardous. Significant effects (p<0.05) are shown in bold. Degrees of freedom are Greenhouse-Geiser corrected when Mauchly's test showed violation of sphericity

Effect	NASA-TLX				RSME			
	df 1,2	F	p	η_p^2	df 1,2	F	p	η_p^2
MAT	**1,48**	**160.214**	**<0.001**	**0.769**	**1,48**	**154.737**	**<0.001**	**0.763**
Group (G)	1,48	2.072	0.157	0.041	**1,48**	**5.773**	**0.020**	**0.107**
MAT x G	1,48	0.286	0.595	0.006	1,48	0.566	0.455	0.012
Traffic Condition (TC)	**3.4, 162.2**	**31.262**	**<0.001**	**0.394**	**4,45**	**11.833**	**<0.001**	**0.513**
TC x G	3.4, 162.2	1.870	0.129	0.037	4,45	0.399	0.808	0.034
MAT x TC	3.3, 159.1	1.248	0.294	0.025	4,45	0.451	0.771	0.039
MAT x TC x G	3.3, 159.1	1.349	0.259	0.027	4,45	0.725	0.580	0.061

Table 5.8 shows univariate test results for subjective measurements of NASA-TLX and RSME. As shown in the table, significant main effects were found on the type of secondary tasks for both NASA-TLX and RSME. Significant main effects of traffic condition were also found for NASA-TLX and RSME. Another significant main effect also found between driving experience group of novice and experienced for RSME rating. The interactions between factors were not significant.

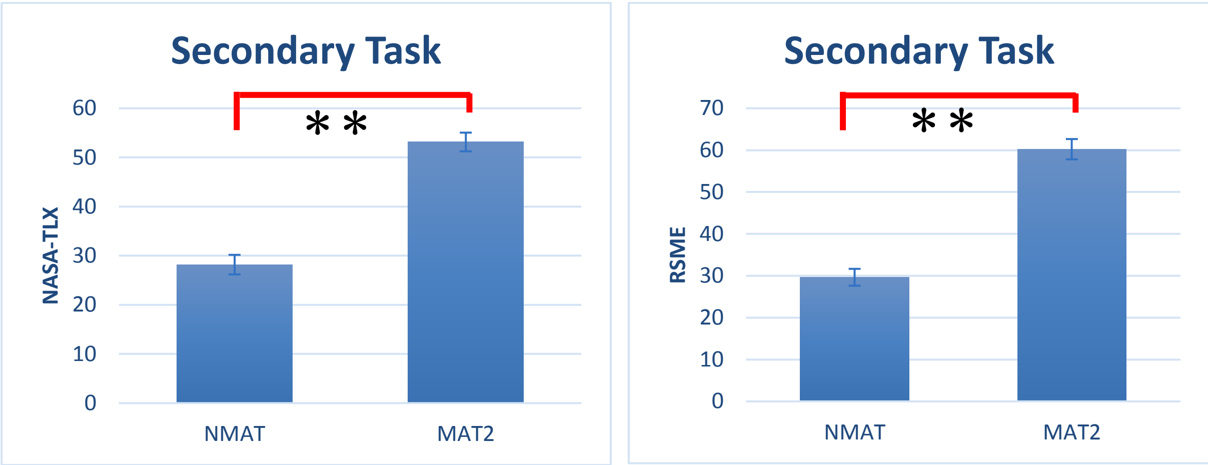


Figure 5.23 Mean value of NASA-TLX and RSME according to type of secondary tasks

Figure 5.23 shows a mean value of NASA-TLX and RSME according to the type of secondary tasks, NMAT and MAT2. From the graphs, participants rated a higher workload for MAT2 than NMAT. This result is expected from the beginning as they were already obtained the results from Experiment#2.

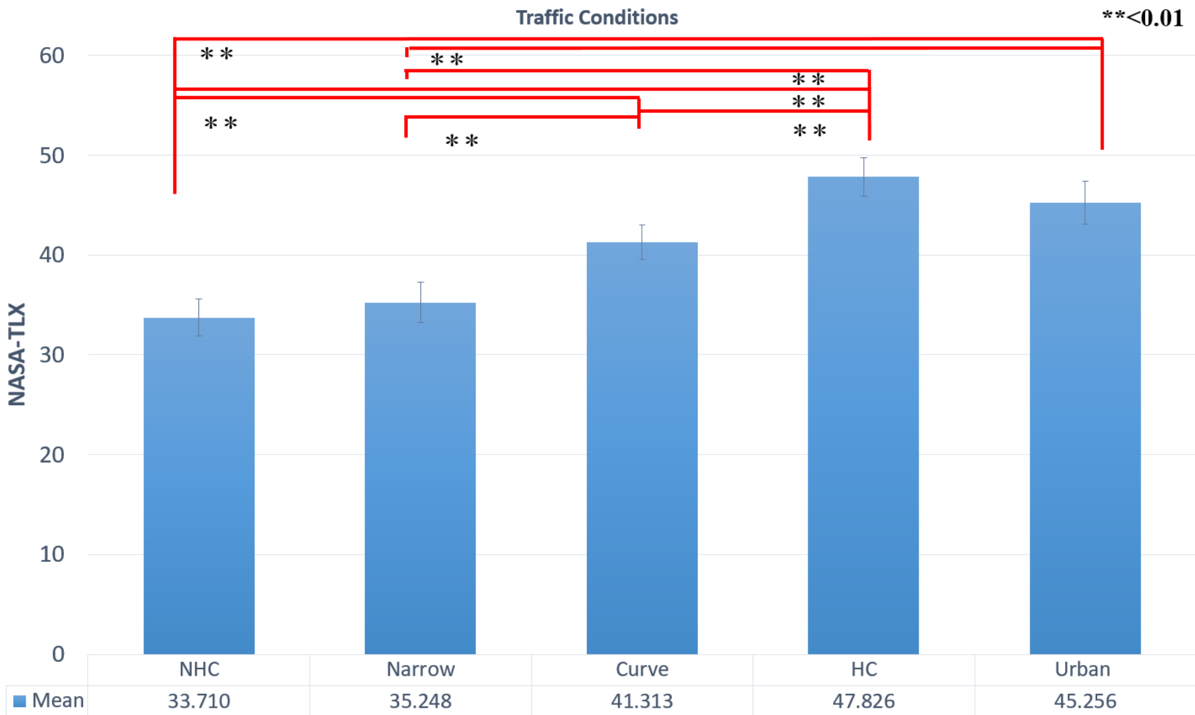


Figure 5.24 Mean value of NASA-TLX according to traffic conditions

Figure 5.24 shows the mean value of rating NASA-TLX for traffic conditions. Participants rated the hazardous condition as the highest mental workload among others. The lowest rating of NASA-TLX was found on NHC.

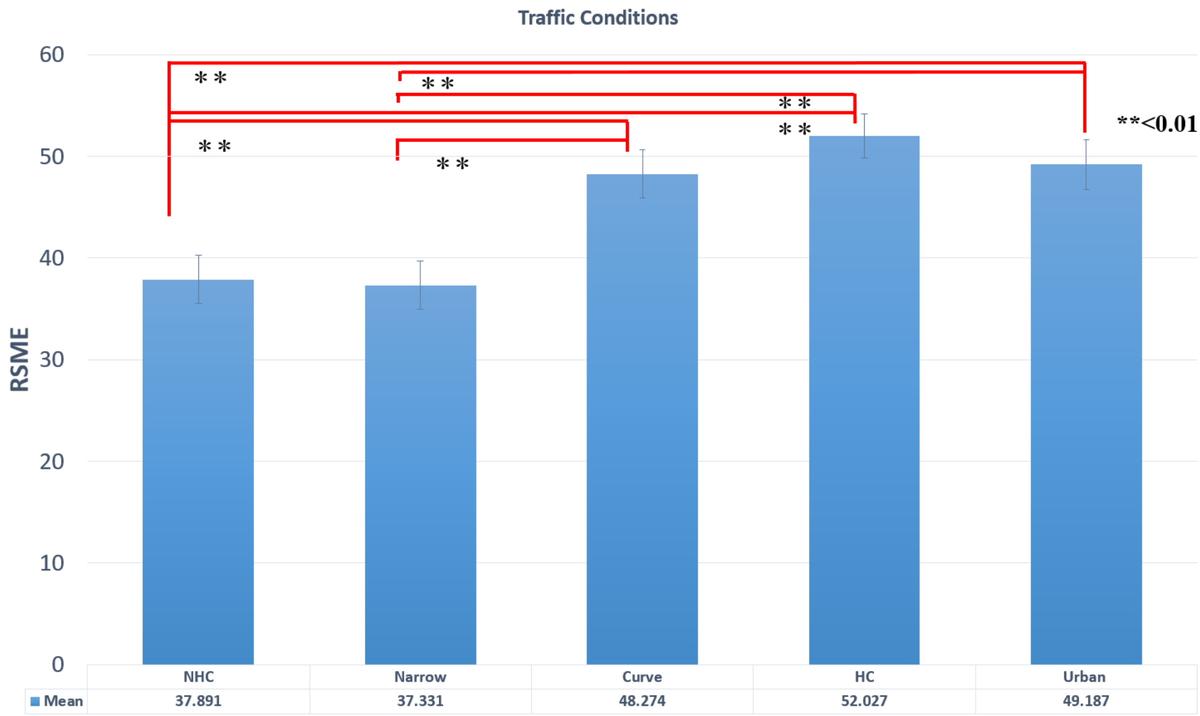


Figure 5.25 Mean value of RSME according to Traffic conditions

Figure 5.25 shows a mean value of RSME according to traffic conditions. The highest rating was found during Hazardous condition and Narrow traffic condition rated as the lowest between other conditions.

5.5 Discussion

5.5.1 The main effects of driving conditions on measurements

As our objective of this experiment was to focus on the external factors of mental workload, we reduce the type of secondary task into two types, either no mathematical arithmetic task (NMAT) or mathematical arithmetic task with a difficulty level (MAT2). In this experiment, we only chose the difficult level of MAT which we found gave effects on mental workload based on our previous experiment. The results of this experiment were consistent with our expectation that significant main effects of MAT were found on all measurement groups. To discuss further, we want to get an answer on under variation of traffic condition which one among the factors gives effects more than another.

Table 5.9 Summary of significant main effects of driving conditions on the measurements

		MAT (p)	η_p^2	Traffic Condition (p)	η_p^2	Traffic Condition x MAT (p)	η_p^2
Physiological	HR	<0.001	0.482	0.437	0.024	0.103	0.004
	Normalized HR	<0.001	0.513	0.339	0.008	0.032	0.020
	HRV	0.005	0.054	0.126	0.013	0.405	0.007
	Normalized HRV	0.530	0.003	0.170	0.012	0.212	0.011
Performance	SDLP	<0.001	0.678	<0.001	0.768	0.001	0.109
	Normalized SDLP	<0.001	0.711	<0.001	0.738	0.004	0.084
	SE	<0.001	0.763	<0.001	0.865	0.002	0.307
	Normalized SE	<0.001	0.757	<0.001	0.862	0.002	0.312
	Percentage of Correct answer	X	X	0.157	0.161	X	X
Subjective	NASA-TLX	<0.001	0.769	<0.001	0.394	0.294	0.025
	RSME	<0.001	0.763	<0.001	0.513	0.771	0.039

As our objective of this experiment was to focus on the external factors of mental workload, we reduce the type of secondary task into two types, either no mathematical arithmetic task (NMAT) or mathematical arithmetic task with a difficulty level (MAT2). In this experiment, we only chose the difficult level of MAT which we found gave effects on mental workload based on our previous experiment. The results of this experiment were consistent with our expectation that significant main effects of MAT were found on all measurement groups. To discuss further, we want to get an answer on under variation of traffic condition which one among the factors gives effects more than another.

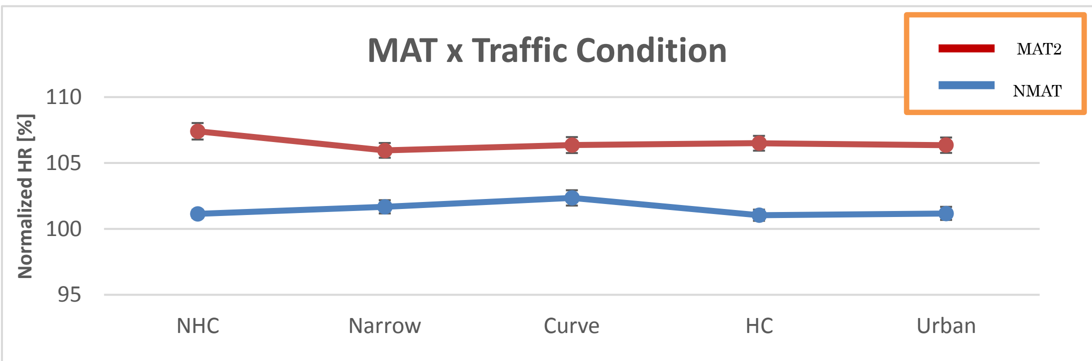
Table 5.9 shows a summary result of significant main effects based on regroup-measures ANOVA to each measurement groups. A yellow color of the column represents a significant main effect based on p-value less than 0.5. While the blue color of the column indicates the value of effect size of partial eta square which was dominant among other effects (MAT, Traffic Condition, and interaction between Traffic condition and MAT) in the measurement.

For the traffic conditions, significant differences were found in Performance measurement groups and Subjective measurement groups. While for physiological groups, no significant were found on traffic conditions. The value of effects size of MAT was higher on physiological measurements and subjective measurements which indicate that these measurements were sensitive to measure internal factors of mental workload such as cognitive workload. Other than that, performance measurements are more sensitive if the experiment involves the manipulation of external factors of mental workload. Although performance measurements were also sensitive to assess mental workload on internal factors, e.g. SDLP was reduced when participants accomplishing mental tasks, based on our experiment, traffic condition give more effect to participants.

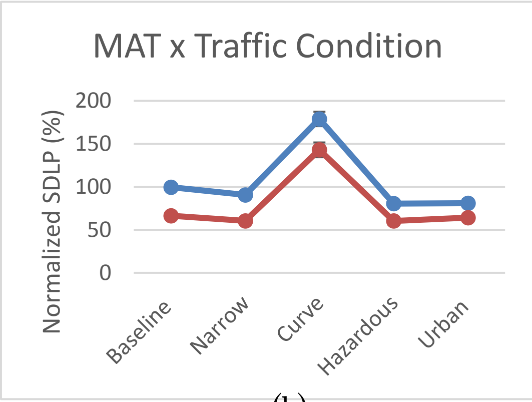
These results also suggest that if we want to measure the mental workload between different traffic conditions, the performance measurement group will be a good measurement to be selected because it is directly related to driving performance. For physiological measurements, at least in this experiment as we used HR as physiological measurements, did not found to be sensitive to the change of traffic condition designed in this experiment. Participants might not feel any differences based on physiological changes in different traffic conditions, but when we look at the performance, their performance deteriorated.

5.5.2 The interaction between internal and external factors of mental workload

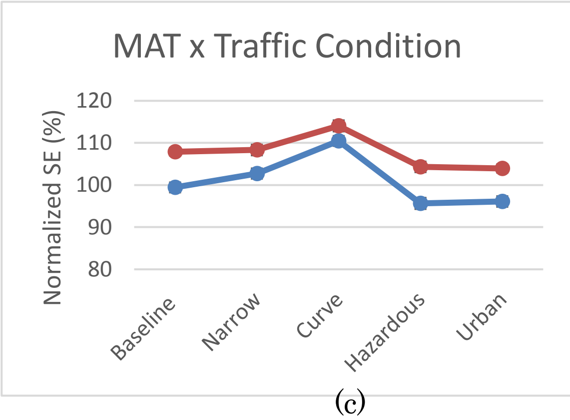
To answer the question on whether the combination of internal and external factors will further increase the mental workload, we examine the significant interaction between MAT and Traffic Condition. Based on our results, the interaction between MAT and Traffic conditions were found significant in Physiological measurements: Normalized HR, and Performance Measurements: SDLP, Normalized SDLP, Steering Entropy and Normalized Steering Entropy.



(a)



(b)



(c)

Figure 5.26 Interaction between MAT and Traffic Condition

Figure 5.26 shows the interaction graphs between MAT and Traffic condition. The results were consistent with our finding in Experiment#2, accomplishing the higher level of the mental task will increase the mental workload. Based on the results of physiological and performance measurements, we conclude that accomplishing the mental task will further increase mental workload and deteriorating driving performance. The different results between Normalized SDLP and Steering Entropy when accomplishing MAT2 or Not. While accomplishing MAT2, the SDLP was lower than during NMAT and Steering Entropy was increased. When looking closely this results is consistent with Experiment#2 which we found that internal factors will increase the Steering Entropy and decrease SDLP. We think that as the nature of Steering Entropy measurement, it is an indicator of smoothness in steering control. These results are showing that, although the SDLP was reduced and seems to improve the performance, however, the steering control of the vehicle under this situation was increased. Maybe Steering Entropy is more sensitive to detect the effect of internal sources better than SDLP.

5.5.3 Relationship between Mental Workload and Driving Performance

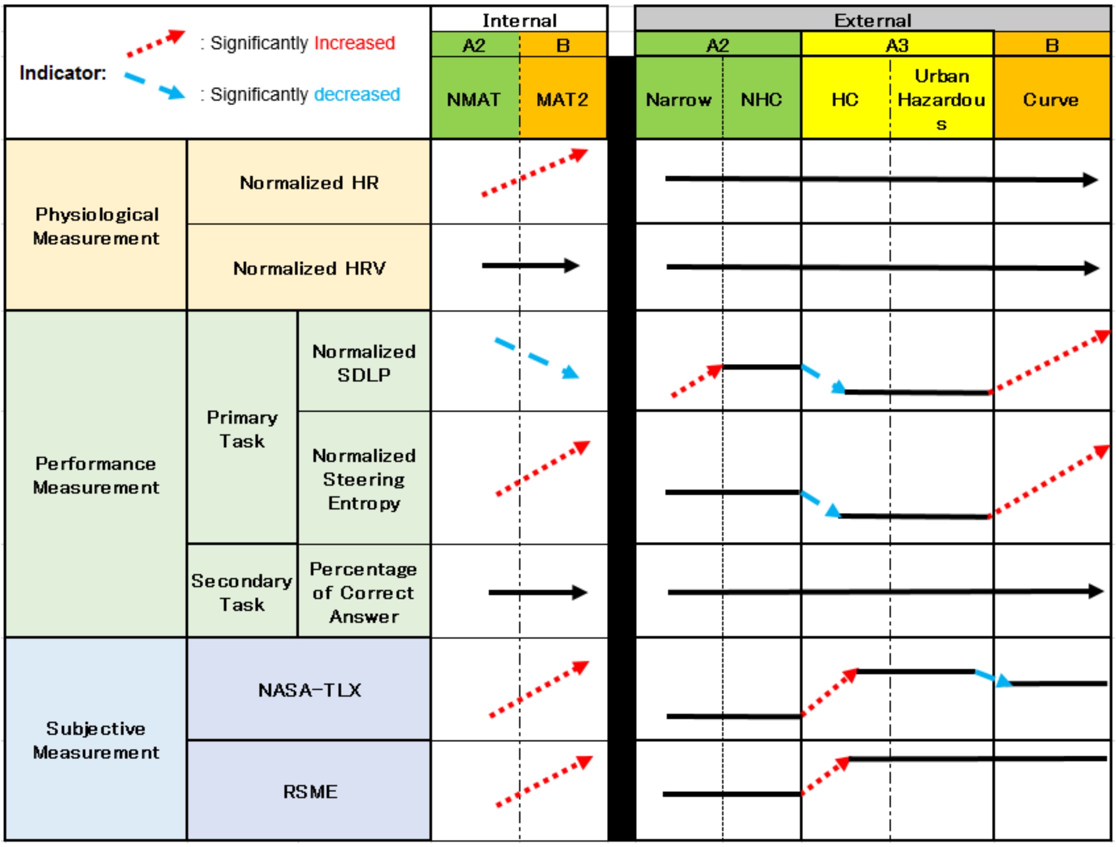


Figure 5.27 Overall results for evaluation of mental workload

Figure 5.27 shows the overall results for evaluation of mental workload. From the results of the experiment and considering the significant effect among traffic conditions we tried to map into a certain region based on de Waard's model [20].

The most obvious finding from this experiment was when we evaluate the lateral position performance measurements, which were SDLP and Steering Entropy, the results show that the continuous curve traffic condition is undoubtedly demanding a higher effort. Driving in this kind of traffic conditions will increase the level of mental workload. In terms of safety, participants were swerving more under this traffic condition, and this is a dangerous situation where they could meet with an accident with another vehicle or oncoming traffic when not able to maintain in cruising lane [87].

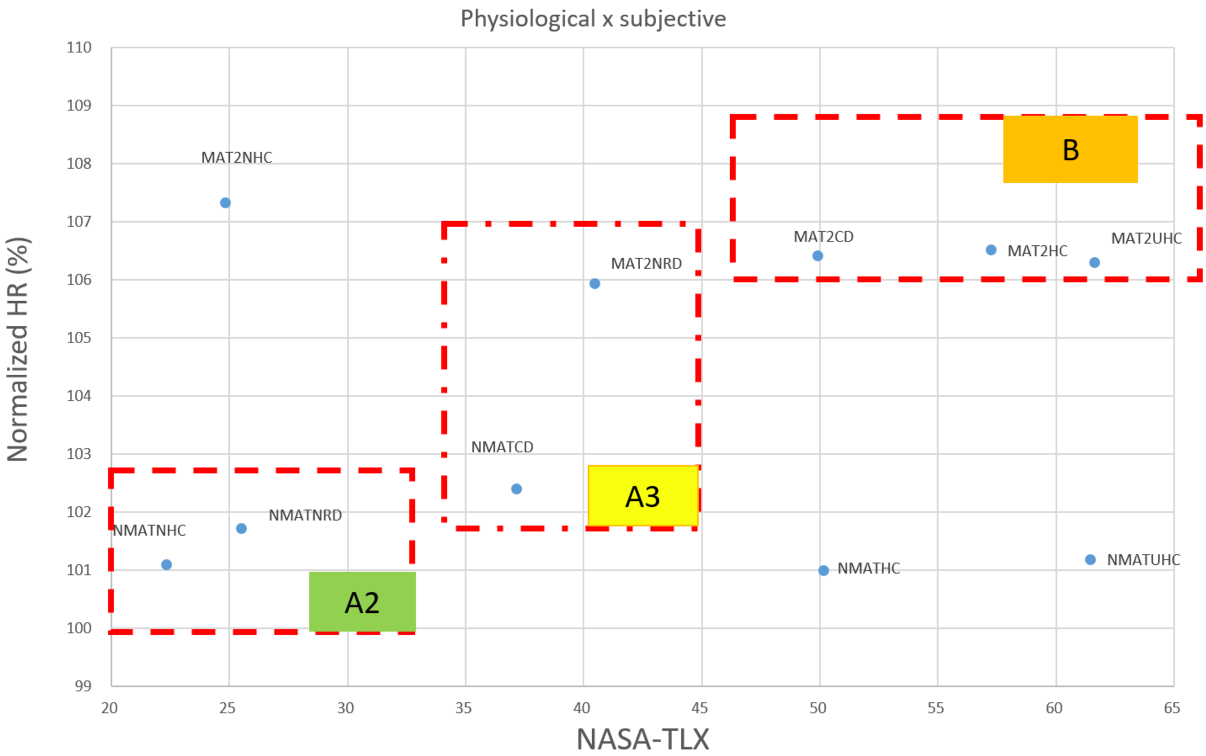
Also to mention here that the result of RSME showed a higher perceived workload when driving in the curved situation and rated as the same with the hazardous condition. In this traffic condition, they need to exert more effort and pay good attention to the Lead Vehicle movements. Considering only for traffic conditions, we think that it is appropriate to map this curve traffic condition into Region B where they are actually in overload. We also think that curve conditions were the most difficult task for the participants. This finding also a strong point in our study as we were able to compare the mental workload measurement between traffic conditions systematically.

Another important finding from this experiment is when participants are driving under a reduced width of the road, the performance was at optimum level as the value of the SDLP was the lowest among others. Moreover, the result from normalized SDLP shows that driving on the narrow road was significantly below the baseline driving. The results of Steering Entropy under this traffic condition was not much difference with NHC, mean that they were trying their hard to control the steering as smooth as they could as in the NHC. SDLP results showed that the performance was better and less swerving than in wider road of NHC. For these reasons, we think it is appropriate to map the Narrow Road condition (NRD) to Region A2 where participants were at the optimum level of performance, and mental workload was the lowest. While for NHC, as they swerved more than in the narrow road, we think it is appropriate to map NHC to A3 Region when they tried their hardest to maintain the performance.

Another point to mention here is the difference between narrow road and baseline driving condition. From the results, we can confirm that participants were actually can distinguish between the NHC and the narrow road. One of the interpretations of this result is training strategy that we adopted in this experiment is effective.

For the Hazardous Condition and Urban Hazardous condition, all the measurements showed that there was no difference between these two traffic conditions. We designed the two conditions were totally the same traffic condition, and the difference was only for the traffic environment of highway road or urban road. We concluded that driving in urban or highway road will not give a different if the movement of other objects not related to their point of attention. Maybe by adding some of the complex situations such as the existing of traffic lights, driving on a junction and overtaking will give further interesting findings on this point. For these two traffic conditions, based on the results, participants were tried their hardest to maintain performance, and it is appropriate to map these traffic conditions to Region A3.

As for internal factors, the results of this experiment were consistent with the results in Experiment#2.



*NMAT: No Mathematical Arithmetic Task, MAT2: Mathematical Arithmetic Task with two digit number, NHC: Non-Hazardous Condition, NRD: Narrow Road Driving, CD: Curve Driving, HC: Hazardous Condition, UHC: Urban Hazardous Condition

Figure 5.28 Mapping between Normalized HR and NASA-TLX

Figure 5.28 shows a mapping graph between a physiological measurement of Normalized HR and a subjective measurement of NASA-TLX. As shown in the mapping graph, under NMATNHC and NMATNRD conditions where participants did not perform any MAT, Normalized HR was almost 100% and close to the value of Baseline Driving (100%). While for NASA-TLX, participants rate both NMATNHC and NMATNRD were the lowest among others. By considering these points, it is appropriate to map these two conditions under region A2 where the drivers were in low workload.

While in conditions of MAT2CD, MAT2HC, and MAT2UHC, the Normalized HR and NASA-TLX were about in the same location. It is suitable to map these two conditions into Region B, they actually in overload region.

For MAT2NRD and NMATCD, they were still in high mental workload, and it is appropriate to map these two driving conditions into A2, where participants were exerting their effort to cope with the increasing demand of workload.

We found through this experiment; several driving conditions did not match with perceived workload through physiological measurement of HR and subjective rating of NASA-TLX. For example, under all conditions of MAT2, where participants accomplished a difficult level of secondary task, a higher Normalized HR can be observed. However, when coming to the rating of how much load did they perceived, they rated differently between traffic conditions where MAT2NHC was among the lowest. Conversely, for NMATHC while they are accomplishing only for driving under the urban hazardous condition, the rating was among the highest while the Normalized HR was among the lowest. This phenomenon, in some conditions, can be explained as dissociation between subjective and physiological measurements. This dissociation between measurements has been described by some of the researchers such as [15], [83]–[86], however, in this experiment, we can validate the dissociation by manipulating experimental conditions between internal and external factors of mental workload. As the perceived workload of subjective measurement also depends on the traffic situation. Our strong point to present here is we were able to map between these two measurements by normalizing the physiological measurements.

5.5.4 Driving Experience

Another contributing factor of increasing mental workload is driving experience. Through the experience of driving the way, the novice drivers and experienced drivers perceived the workload is different. Experienced drivers are said to automate some of the driving tasks, making them better in processing more workload resulting better in driving performance.

In this experiment, regarding the differences between Novice group and Experienced Group, we could see the differences in some of the measurements. First, we can observe through the subjective measurements of RSME.

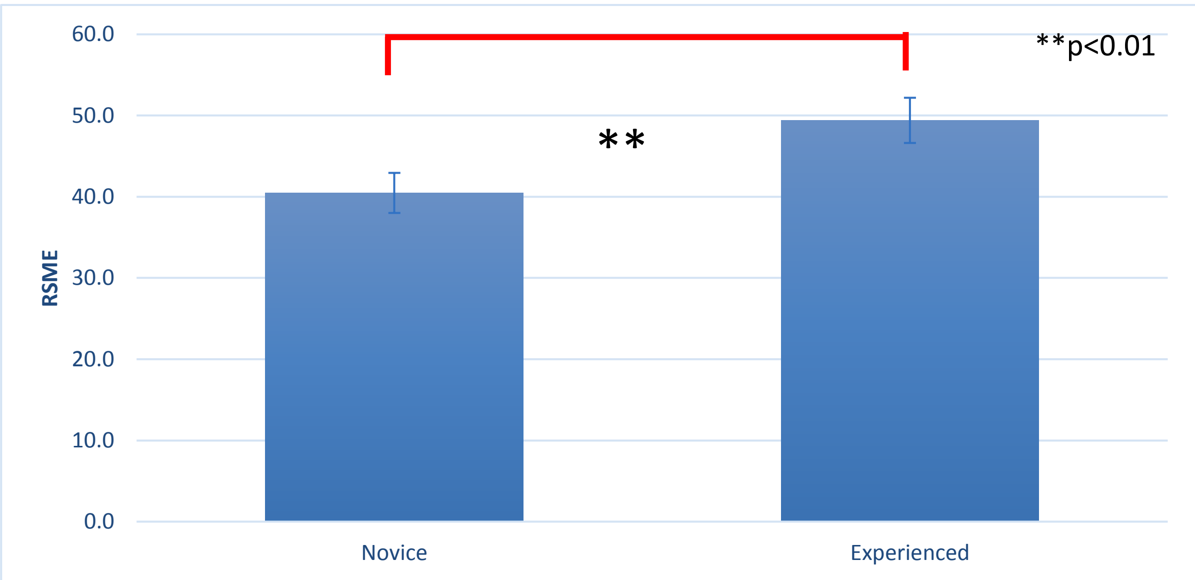


Figure 5.29 Differences between Novice and Experienced Drivers in RSME

Figure 5.29 shows a graph from the results of a rating scale of mental effort (RSME) between novice and experienced drivers. Contradictory to our expectation, experienced drivers rated overall task in our experiment higher than novice drivers. In spite of the fact that in this experiment, we focused on traffic conditions which are experienced drivers expert in, we think that the secondary task of mental workload affects the way they perceived the workload. This point can be confirmed by looking at the secondary task performance, represented by the percentage of the correct answer which the performance of experienced drivers was significantly worse than novice drivers. This finding is actually consistent with the previous finding on a secondary task which was experienced drivers performed worse than novice drivers in the secondary task [72][41][88]. Although some of these studies used a different kind of secondary task, e.g. Reaction time, but the results show the same trend of reducing the performance of secondary task compared to the younger age of drivers. We think that the way they allocate the priority to the task are the keywords we can use to explain this notation. We can explain in detail about this when looking at the differences between novices and experienced on performance measurements.

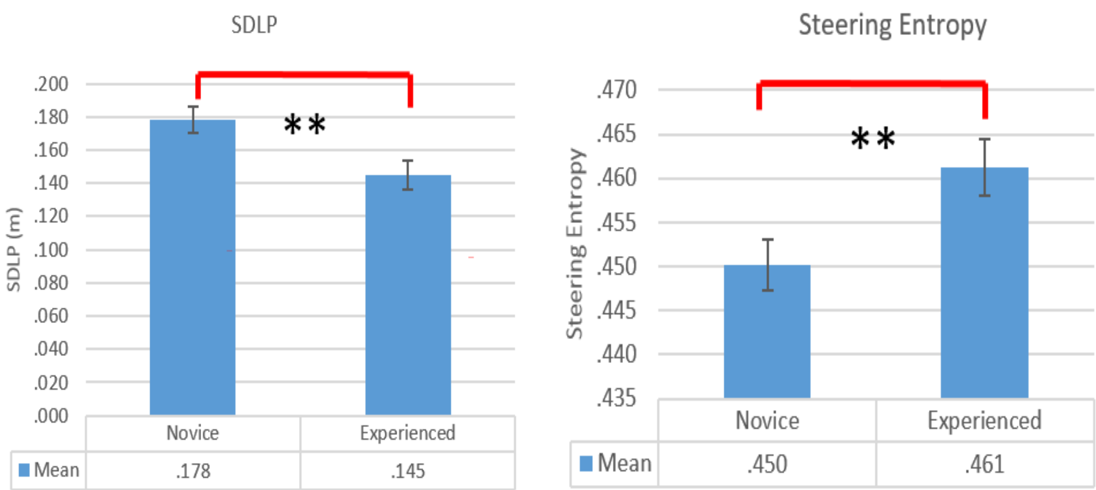
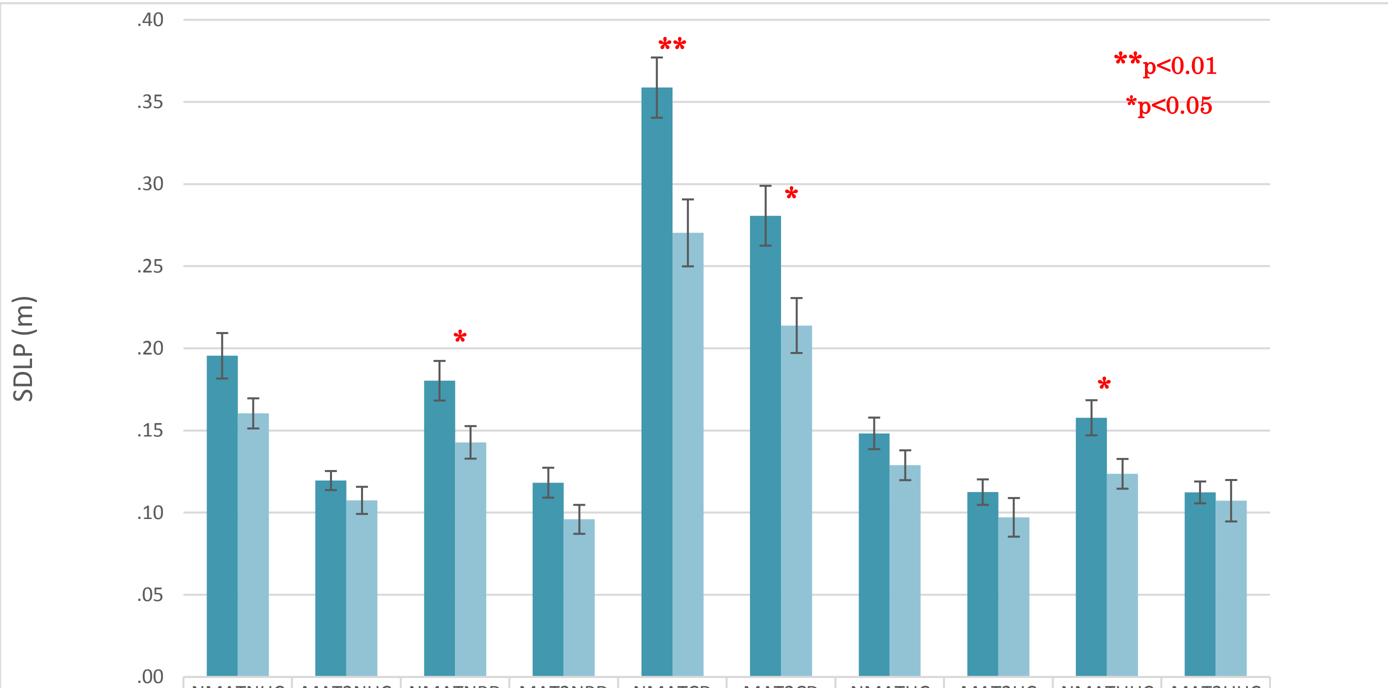


Figure 5.30 SDLP and Steering Entropy results according to novice and experienced group

Figure 5.30 shows results of the SDLP and Steering Entropy according to novice and experienced group. Significant differences were further found between these two groups based on the SDLP and Steering Entropy results.

The result of SDLP shows that the novice driver swerved more than experienced drivers. Interestingly, results of steering entropy which was to measure the smoothness of steering control indicate that the experienced driver gives a higher value of Steering Entropy than novice drivers. This results maybe can be explained as during driving, experienced driver performed better in terms of maintaining the vehicle in the cruising lane than novice drivers. This notation may be resulted from putting the effort in adjusting the steering wheel. Based on this results of performance and putting together the results of secondary task performance, we can synthesize that experienced drivers allocate their priority in controlling the vehicle safely (based on SDLP result) resulting in the performance on secondary tasks deteriorated.

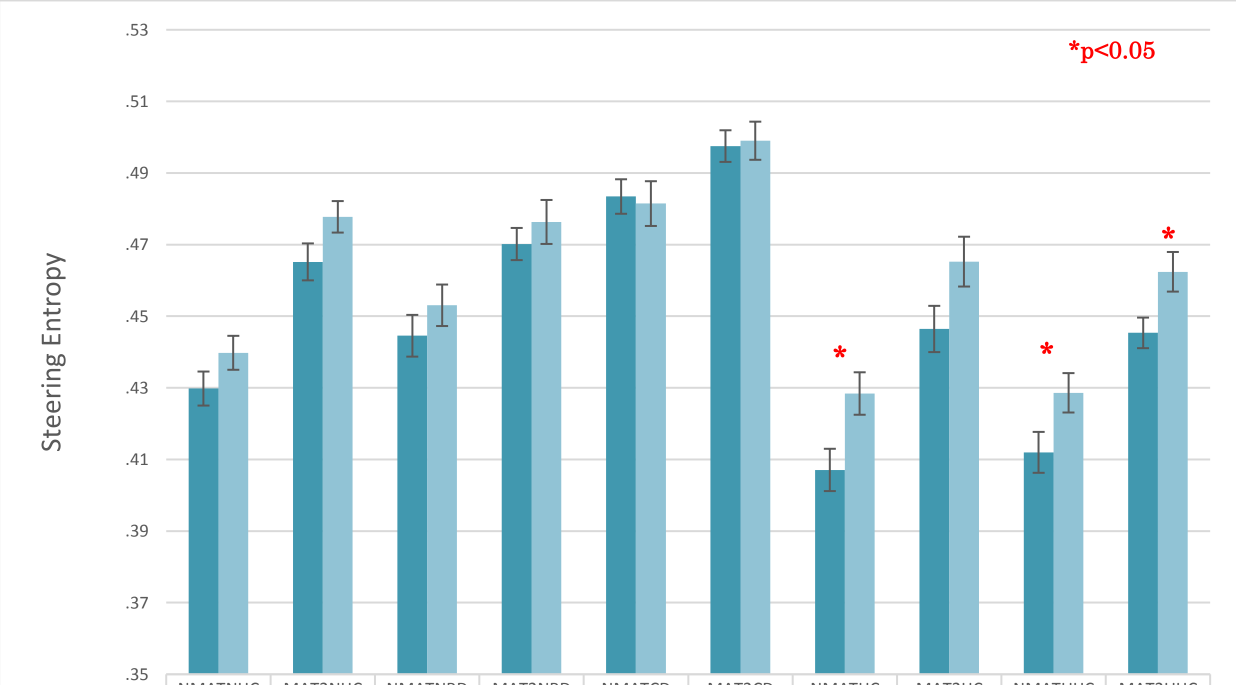
Another point to be mentioned here regarding the reverse result between SDLP and Steering Entropy is the skill of handling a vehicle in terms of maintaining the vehicle within the cruising lane. We further analyzed the differences between these two groups by using independent t-test with driving conditions (NMAT: No MAT, MAT2: Mathematical Arithmetic Task with second digit numbers, No-Hazardous Condition: NHC, NRD: narrow road driving, CD: Curve Driving, HC: Hazardous Condition, and UHC: Urban Hazardous Condition) as independent variables. As shown in Figure 5.31, in all conditions, experienced drivers show a less value of SDLP compared to novice drivers. The differences between novice and experienced groups were found on NMATNRD, t(48)=2.314, p=0.025, on NMATCD, t(48)=3.219, p=0.002, on MAT2CD, t(48)=2.635, p=0.011, and on NMATUHC, t(48)=2.353, p=0.023 respectively. Particularly on curves, experienced drivers show a less swerving compared to novice drivers showing that the skill of handling the vehicle makes the differences.



	NMATNHC	MAT2NHC	NMATNRD	MAT2NRD	NMATCD	MAT2CD	NMATHC	MAT2HC	NMATUHC	MAT2UHC
Novice	.1955	.1195	.1803	.1182	.3587	.2807	.1483	.1125	.1578	.1123
Experienced	.1604	.1075	.1428	.0959	.2703	.2139	.1289	.0971	.1236	.1073

Figure 5.31 Result of SDLP for Novice and Experienced

Figure 5.32 shows a result of Steering Entropy between novice and experienced drivers. Except for NMATCD, the value of Steering Entropy for novice drivers was less than experienced drivers in all driving conditions. The differences between novice and experienced groups were found on NMATHC, t(48)=-2.515, p=0.015, on NMTUHC, t(48)=-2.056, p=0.045, and on MAT2UHC, t(48)=2.481, p=0.017. These results also to confirmed that the higher value of Steering Entropy among experienced drivers was not due to the difficulties in handling the vehicle on curve condition, but maybe from other factors such as traffic complexity in hazardous conditions. Further research needs to be carried out to confirm this point.



	NMATNHC	MAT2NHC	NMATNRD	MAT2NRD	NMATCD	MAT2CD	NMATHC	MAT2HC	NMATUHC	MAT2UHC
Novice	.4298	.4652	.4445	.4702	.4834	.4975	.4071	.4464	.4120	.4453
Experienced	.4398	.4778	.4531	.4763	.4815	.4990	.4284	.4652	.4286	.4624

Figure 5.32 Steering Entropy Results for Novice and Experienced Drivers

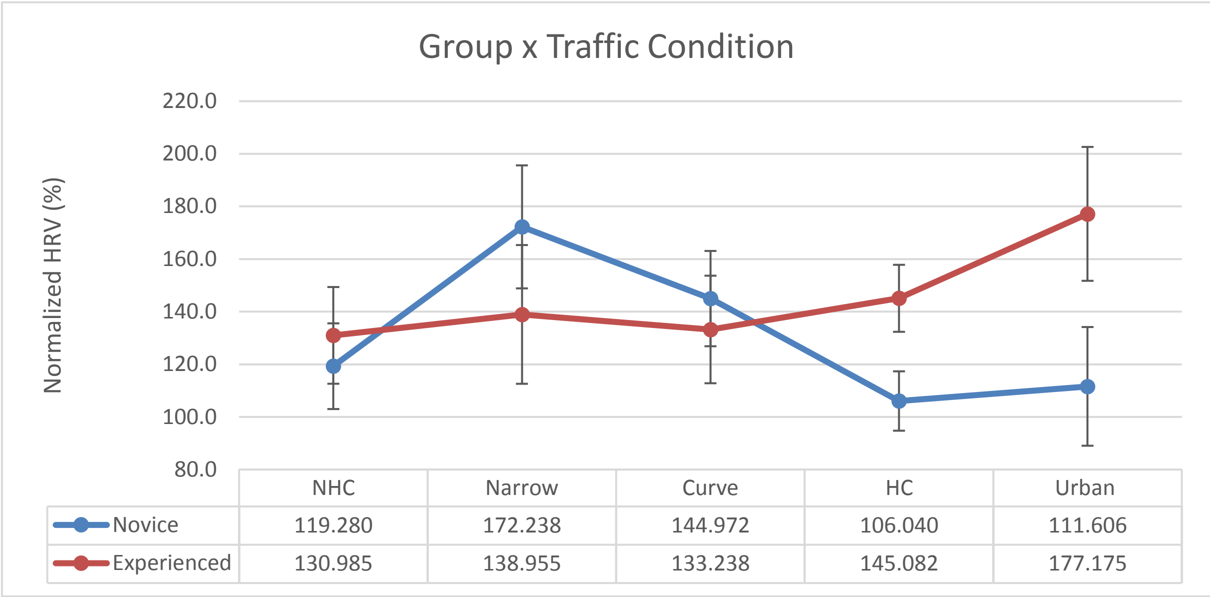


Figure 5.33 Interaction between Group and Traffic Condition

Figure 5.33 shows a graph illustrating the interaction between groups on normalized HRV. As shown in the graph, significant difference between Novice and experienced drivers was found on two traffic conditions. When driving on HC and Urban HC, novice effected more than experienced. This result also an evidence of how experience drivers handle with a complex situation while novice drivers feel difficult in handling the situation.

<u>Conclusion for driving experience</u>

From above results, we concluded that the experienced drivers give priority on the driving to maintain the performance. It can be observed from the decrease in secondary task performance and less swerving especially in curves driving condition. They were able to maintain the performance, and this result the secondary task performance dropped.

While the novice drivers, maybe undermine the safety of driving task, when they exposed to difficult, hazardous conditions. This notation can be confirmed by the lower value for Steering Entropy than experienced drivers in both hazardous conditions. We think that less value in steering entropy means that they made less steering correction compared to experienced drivers. Novice shares most of their focus to the secondary task while for experienced drivers able to maintain driving as their main focus.

5.5.5 Factors of increasing mental workload

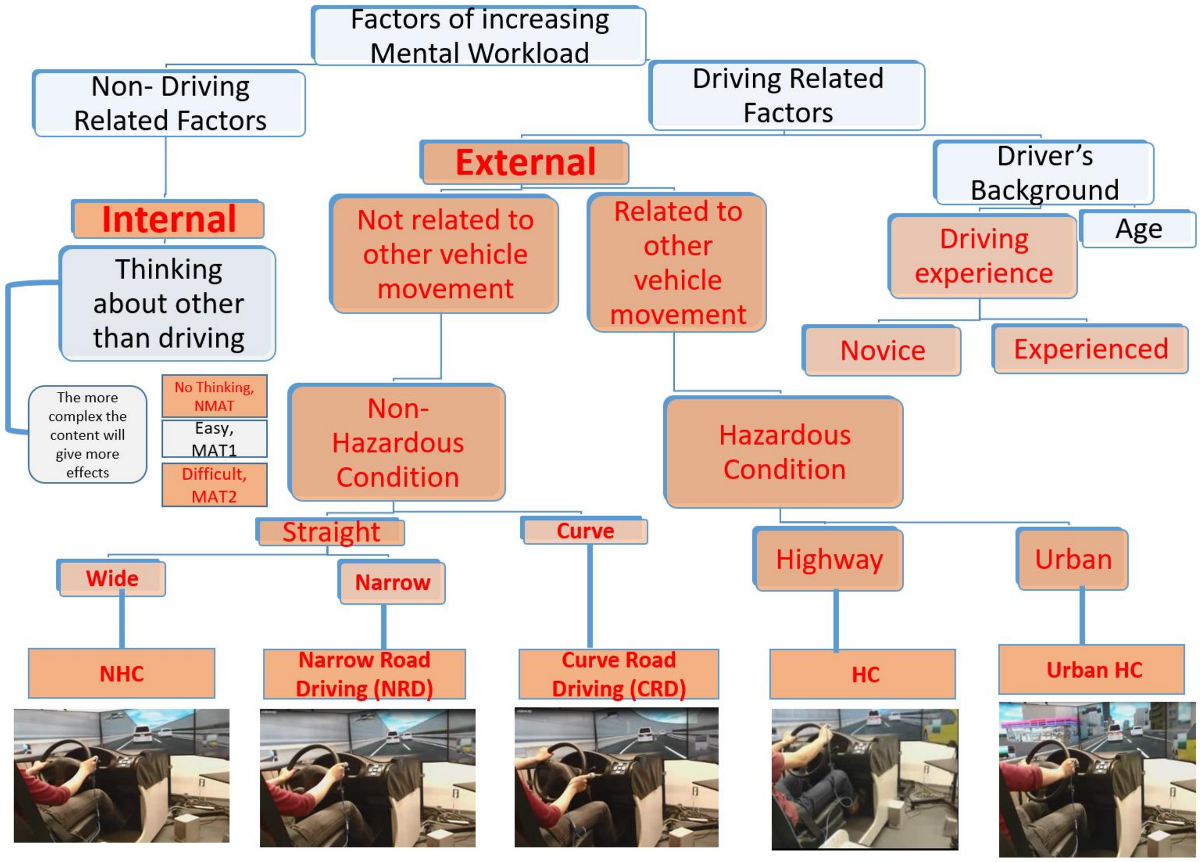


Figure 5.34 Increasing factors in Experiment#3

Figure 5.34 shows a diagram on increasing factors of mental workload in Experiment#3. In this experiment, we focus on the external factors of increasing mental workload. Based on the overall results in Figure 5.27, the differences between Related to other vehicle movement and not related to other vehicle movement were found when we compared between narrow and NC to Hazardous condition or Urban Hazardous condition. There was no significant difference between a hazardous condition and urban hazardous condition in all measurement groups. This result to confirm that although we change the environment by driving from the highway to an urban, there is no effect on the measurements under different environment. The differences were found between Narrow road and NHC when compared to a hazardous condition on steering entropy. The results show that driving under hazardous condition will make both SDLP and Steering Entropy decreased. It can be understood as when they are driving under a risky situation that needs to exert their attention more on the driving situation, they control the steering less and less swerving.

5.6 Conclusion

This experiment was designed to investigate external factors of mental workload. Five traffic conditions where we think will influence the increasing mental workload been designed for this experiment. Namely, Non-Hazardous Condition, Narrow Road Driving, Curve Driving, Hazardous Condition and Urban Hazardous Condition.

These are the conclusions of this experiment:

1. MAT gives effects on to the drivers on physiological measurements and subjective measurements.
2. Traffic Condition gives more effect on Performance measurements
3. Depending on the driving conditions, performance could be degraded. Obviously appeared in curve driving condition.
4. Based on the traffic conditions, driving with higher difficulties in internal factors under a complex traffic situation of external factors will further increase mental workload and performance may become worse.
5. Experience in driving can improve the way a driver perceive workload from internal or external factors.

Chapter 6 General Discussions and Conclusion

6.1 General Discussion

6.1.1 Increasing the accuracy of evaluation for mental workload

In this study, we tried to evaluate the level of the mental workload from internal and external of mental workload. There are two ways that we think will increase the accuracy of evaluation of data and normalization of data.

One of the problems with the data was there was no standard value of mental workload that we could use to evaluate the driver's mental workload. With the individual differences between participants and the effect of "motivation", "mood" of participants, would further increase the uncertainty of data. One of the solutions to this problem was by normalizing the data. In the first experiment, we tried to evaluate the selection of based interval will increase the accuracy of data. By bringing data of physiological measurements, we normalized it by using two based intervals, namely Normalized 45s and Normalized 120s.

In the first experiment, we tried to find the effects of combination level of difficulties from internal and external factors of mental workload.

First, we attempt to evaluate the normalization period by selecting the shorter time of 45s or a longer time of 120s while performing the driving task. The evaluation of normalization 45s and 120s, however, did not find any significant differences between selecting these two periods. To discuss further this, we tried to perform a statistical analysis between driving conditions, for at least to find if there were significant between the driving conditions, mainly to evaluate the level of difficulty of the tasks. As our hypotheses that, the combination of internal and external factors of the mental workload would further increase the mental workload.

Based on the results of significant differences between traffic conditions, we were at least able to show a significant difference in two conditions based on data of normalized 120s, when compared to non-normalizing it which was significant differences between baseline driving on the last day and MAT1NHC and also in MAT2NHC. This result is important for our conclusion in Experiment#1. We were able to distinguish between the day of participants performing the tasks and during the baseline driving for the first and the last day of experiments.

We used the knowledge of normalizing data using 120s that we attained from Experiment#1 to further investigate about internal and external factors of mental workload in Experiment#1 and Experiment#2. In Experiment#1, the driving course was designed to have curves and this we think maybe as a factor of increasing value of mental workload. So, it was not appropriate to use the based interval where the drivers actually in a high mental workload due to one of the external factors in increasing mental workload.

In Experiment#2, as we designed only a straight pathway, we adapt the normalization period of 120s before the MAT interval. Based on the results of physiological measurements, we think that we are able to predict the mental workload by normalizing the value and selecting 120s in a straight pathway as a based interval. We used the same knowledge to evaluate the data of Experiment#3 and using a straight pathway, 120 seconds before MAT interval.

The second thing that we did for increasing the accuracy in the interpretation data on the mental workload level is by comparison between groups. As we have discussed before, there was three measurements group that we can use to evaluate the mental workload level. All of the experiments conducted in this experiment considered the combination of measurement groups to evaluate the mental workload better. For example, in Experiment#1, we combined between Normalized Lyapunov exponent 120s with NASA-TLX to evaluate the level of mental workload. By doing this way, although we knew that MAT2HC (the most difficult level) was in high mental workload. However, cannot differentiate between levels of tasks. However, in Experiment#2 and Experiment#3, we able to differentiate between the levels according to driving conditions.

6.1.2 Evaluation of level of difficulties through combination of Internal and external factors of mental workload

In this study, we attempted to evaluate the level of difficulties from internal and external factors of mental workload. Two main questions raised in this experiment regarding the model of mental workload. 1. Is the combination of internal and external factors will further increase the mental workload and 2. Are the factors giving more effects one another? For internal factors of mental workload, we designed two level of difficulties of the secondary task of manipulating either one or two digit number of mathematical arithmetic task. For external factors of mental workload, there were two levels of difficulties designed as to simulate in a Hazardous condition which participants drive in a situation demanding their full attention to traffic conditions in Experiment#2. While for Experiment#3, we further designed traffic conditions which we think can represent as contributing factors of increasing mental workload.

From the results, we concluded that depending on the traffic situations, the effects of either internal or external workload would appear to be dominant one another. To explain this, in Experiment#2 as the traffic condition was easy, the main effect of internal factors represented by a secondary task appeared to be dominant. While in Experiment#3, as the traffic condition demands, especially for the curve, the effect of external factors appeared to dominate.

For the question of whether the combination of the level of difficulties from the factors of the mental workload will further increase the mental workload, based on our results, we concluded that based on the measurements and the type of traffic conditions the combination would further increase the mental workload.

Results interaction between internal and external factors of the mental workload of Experiment#2, RSME, and a Normalized HR shows the evidence on the point of the combination of the two factors will further increase the mental workload. Results from HR and HRV data show that the trade-off between internal and external factors appeared in the situation where the internal factors to be easy level while driving under difficult situation of external factors.

6.1.3 Hypothetical relationship between Performance and Mental workload

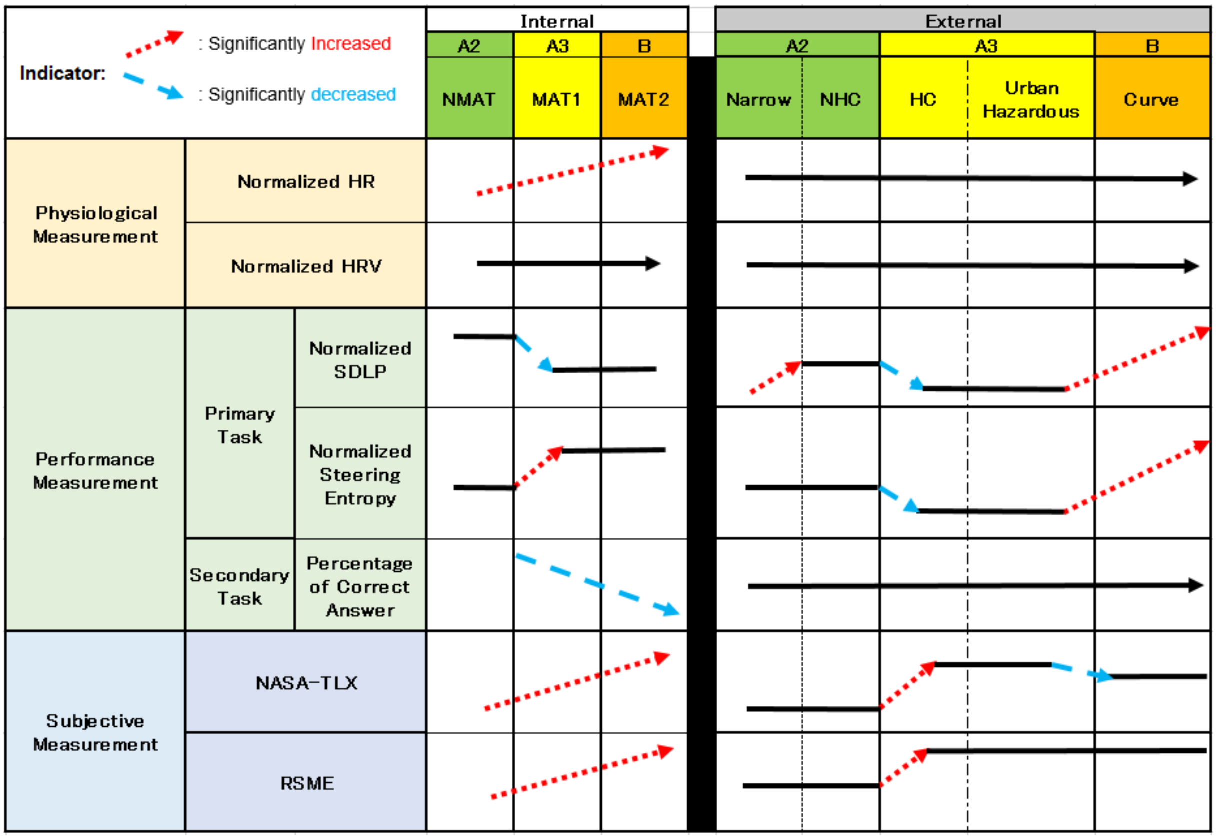


Figure 6.1 Overall results of mental workload measurements based on Internal and External Factors

Figure 6.1 shows overall results of our study based on internal and external factors in Experiment#1-3. As we are mapping together all the measurements, we can get a clearer picture of the relationship between mental workload and performance.

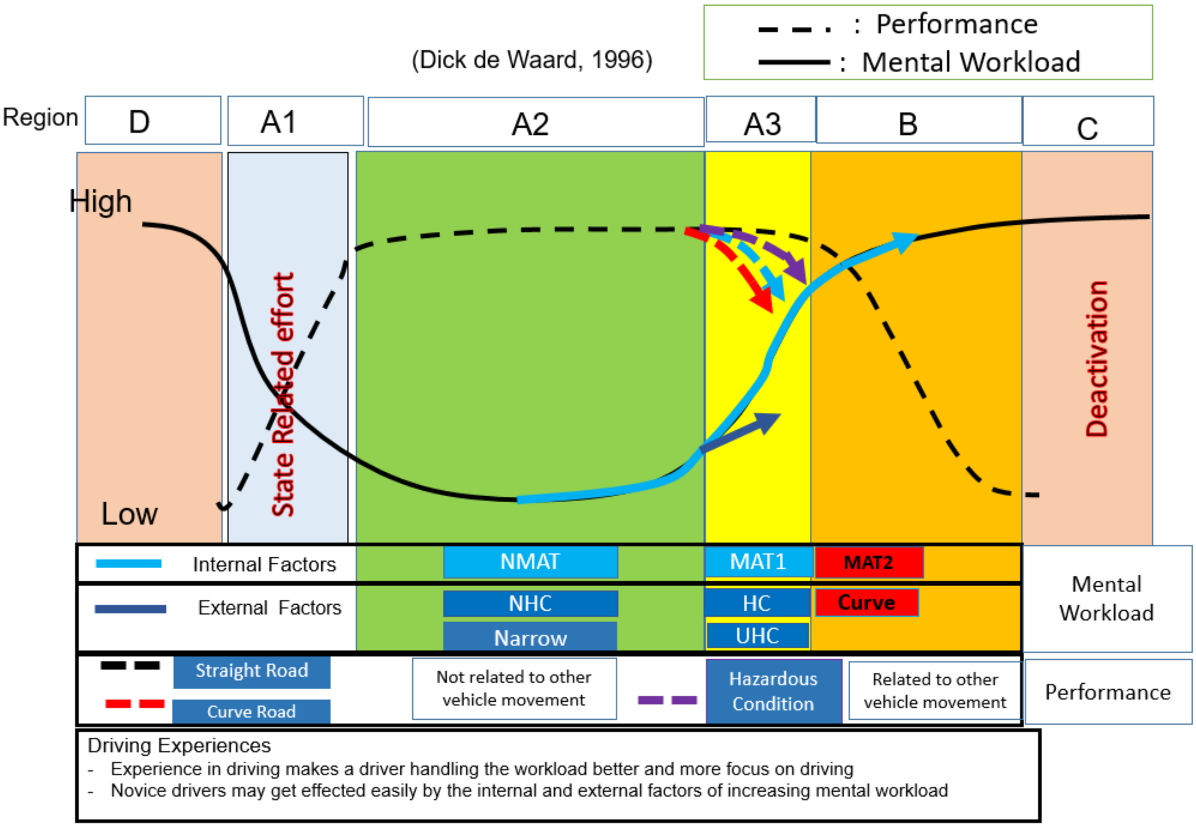


Figure 6.2 Hypothetical relationship between performance and mental workload based on factors

Figure 6.2 shows a hypothetical relationship between performance and mental workload based on the results of our experiments. Overall, we think that we have to think separately the relationship based on from where the factors are: Internal or External. If it is from internal factors, the increasing in mental workload concurrently occurred with the deterioration in performance. External factors are less effect than internal factors for a mental workload. For performance, the deterioration of performance depends on either the factor is related to other vehicle movement or not. If it is related to other vehicle movement, the SDLP and Steering Entropy maybe give a similar trend of decreasing. While if it is not related to other vehicle movement, it depends on the geometry of the road. The curvy road will give drivers in a hard situation to maintain the vehicle within the lane as well as increase the value of steering entropy. We think that maybe Region A3 which Task Related effort does not exist when the drivers are contacting with either internal or external factors. We also think that driving on a curve with the MAT2 condition is the most difficult task in this study and maybe can be regard as Region B where it is overload region. However, further study has to be conducted to validate this hypothesis with more experimental data.

6.2 Conclusion

This thesis describes research undertaken to investigate drivers' mental workload. Three experimental works have been conducted to improve our understanding regarding mental workload.

The main conclusions of this thesis were;

1) The experiments have been able to evaluate the effect of internal and external factors of mental workload to mental workload measurements. The combination of internal and external factors cause the mental workload to be increased.

2) The evaluation of BVP sensor come to the conclusion that BVP sensors could be used in a driving situation as it appears to be non-intrusive to drivers and the data were not much difference with data from the ECG.

3) Since the normalization of data and the combination of measurement groups, we drew maps, graphs between measurement groups, the evaluation of mental workload level appeared to be more systematic and informative. We can see through our study in the discussion of Experiment#1, Experiment#2 and Experiment#3.

4) The internal factors appeared to give more effects to driver's mental workload for physiological and subjective measurements. Performance measurements are sensitive to detect the effect of increasing mental workload to the driving performance.

5) The combination of internal and external factors will further increase the mental workload.

6) Depending on the factors of mental workload, the performance of mental workload appeared to be improved or worse.

6.3 Limitation of study

Although we have achieved our goals in this study, there are some limitations of the study.

1. For the secondary task design, we used only Mathematical Arithmetic Task to represent the difficulties of internal in mental workload. Based on our results of comparison between MAT2NDT in Experiment#3, we cannot confirm whether the participants really think to answer it or not. Maybe we can design another secondary

task of NMAT- repetition of the numbers presented. So at least, they think of the task rather than thinking another thing.

2. Regarding the relationship between lateral performance and mental workload, throughout our experiments, we found that driving while performing the secondary tasks improved the performance. However, we argued that this result may be due to our design of the experiment, e.g. the existing of front vehicle. It is worth investigating further this point.
3. For the selection of normalization data based interval, we tried to select between two based intervals by considering two points 1) Geometry of road 2) Length of time. So, we select 45s and 120s. Based on the results, we found that the difference between these two intervals was not significant. The selection between the intervals may be affected by the geometry or length of time.
4. In this study, we collected data only from male participants due to the limitation in attaching the ECG electrodes to the chest of participants. Maybe for a better estimation of results, we can collect the data from female participants by hiring female research assistants to assist with the procedure.
5. In this study, we only considered the "task-related effort" (de Waard's model refers this as Region A3). However, during the experiment, there might be "state-related effort" such as motivation to perform the task, arousal and another effort which also affects the results.

6.4 Further Research

The followings present some of the issues which require further consideration and possible experimental investigation and which are mainly derived from the limitations of the experimental methods employed in the current work.

1. Further consideration of the types of the secondary task being used in this experiment.
2. Further, study the relationship between lateral performance and mental workload.
3. Further evaluation about the normalization of data should be made.
4. In this experiment, further analysis on gender and the effects of mental workload may be needed in order to make a thorough conclusion regarding these factors.
5. The further study relates to "State Related Effort" is maybe needed.

Acknowledgments

‘In the Name of Allah, the Most Beneficent, the Most Merciful. All the praises and thanks be to Allah, the Lord of the 'Alamin’.

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I am deeply indebted to Prof. Dr. Itoh Makoto whose help, stimulating suggestions and encouragement helped me in all the time of research for and writing of this thesis. I am also thanking Prof. Dick de Waard for his visit to our lab and give meaningful comments and critics with the ideas.

Thanks are extended to the members of the Laboratory for Cognitive Systems Science for their help in developing the idea through discussions and seminars. Especially Ms. Kunugi, who help me a lot in administrative paper works.

Lastly and most importantly, I would like to thank my wonderful wife, W Fatin Zuharah W Musthapa for her patient and limitless support throughout my study. Thanks also to my lovely daughters, Nusayba Khunaisa, Asma Khayrah and Afra Khawlah who made my journey in my study become meaningful and make me strive harder for our future. My greatest thanks go to my parents who raised me up. My family who limitlessly supports me.

Bibliography

[1] Police Bureau of Transportation, “Statistics of Road Accidents in Japan year 2015,” 2016.

[2] B. Reimer, “Impact of Cognitive Task Complexity on Drivers’ Visual Tunneling,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2138, no. 1, pp. 13–19, 2009.

[3] L. Angell, J. Auflick, P. a. Austria, D. Kochhar, L. Tijerina, W. Biever, T. Diptiman, J. Hogsett, and S. Kiger, “Driver workload metrics project: Task 2 final report,” 2006.

[4] A. Sonnleitner, M. S. Treder, M. Simon, S. Willmann, A. Ewald, A. Buchner, and M. Schrauf, “EEG alpha spindles and prolonged brake reaction times during auditory distraction in an on-road driving study.,” *Accid. Anal. Prev.*, vol. 62C, pp. 110–118, Sep. 2013.

[5] H. Makishita and K. Matsunaga, “Differences of drivers’ reaction times according to age and mental workload,” *Accid. Anal. Prev.*, vol. 40, no. 2, pp. 567–575, 2008.

[6] M. M. Haque and S. Washington, “A parametric duration model of the reaction times of drivers distracted by mobile phone conversations.,” *Accid. Anal. Prev.*, vol. 62C, pp. 42–53, Sep. 2013.

[7] W. Consiglio, P. Driscoll, M. Witte, and W. P. Berg, “Effect of cellular telephone conversations and other potential interference on reaction time in a braking response,” *Accid. Anal. Prev.*, vol. 35, no. 4, pp. 495–500, 2003.

[8] K. Brookhuis, D. De Waard, and B. Mulder, “Measuring driving performance by car-following in traffic,” *Ergonomics*, vol. 37, no. 3, pp. 427–434, 1994.

[9] A. J. McKnight and A. S. McKnight, “The effect of cellular phone use upon driver attention,” *Accid. Anal. Prev.*, vol. 25, no. 3, pp. 259–265, 1993.

[10] J. L. Harbluk, Y. I. Noy, P. L. Trbovich, and M. Eizenman, “An on-road assessment of cognitive distraction: Impacts on drivers’ visual behavior and braking performance,” *Accid. Anal. Prev.*, vol. 39, no. 2, pp. 372–379, 2007.

[11] Y. Liang and J. D. Lee, “Combining cognitive and visual distraction: Less than the sum of its parts,” *Accid. Anal. Prev.*, vol. 42, no. 3, pp. 881–890, 2010.

[12] C. J. D. Patten, A. Kircher, J. Ostlund, and L. Nilsson, “Using mobile telephones: cognitive workload and attention resource allocation.,” *Accid. Anal. Prev.*, vol. 36, no. 3, pp. 341–50, May 2004.

[13] C. L. Baldwin and J. T. Coyne, “Mental workload as a function of traffic density: comparison of physiological, behavioral, and subjective indices,” *Second Int. Driv. Symp. Hum. Factors Driver Assessment, Train. Veh. Des.*, pp. 19–24, 2003.

[14] J. Törnros and A. Bolling, “Mobile phone use – effects of conversation on mental workload and driving speed in rural and urban environments,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 9, no. 4, pp. 298–306, 2006.

[15] R. W. Backs, M. Pleasant, J. K. Len-, G. M. Corporation, J. M. Wetzel, and C. Michi-,

"Cardiac Measures of Driver Workload during Simulated Driving with and without Visual Occlusion," *Hum. Factors Journal of Hum. Factors Ergonomics Society*, vol. 45, no. 4, pp. 525–538, 2003.

[16] P. Green, "Driver Workload as a Function of Road Geometry : A Pilot Experiment," University of Michigan Transportation Research Institute, Technical Report UMTRI-93-39, 55 pages, December 1994.

[17] A. Brouwer and C. Dijksterhuis, "Physiological correlates of mental effort as manipulated through lane width during simulated driving," in Proceedings of *International Conference on Affective Computing and Intelligent Interaction (ACII) Physiological*, 2015, pp. 42–48.

[18] J. Michon, "A critical view of driver behavior models: what do we know, what should we do?," *Hum. Behav. Traffic Safety*, pp. 485–520, 1985.

[19] D. Meister, *Behavioral Analysis and Measurement Methods*. New York, New York, USA: John Wiley and Sons, 1985.

[20] D. de Waard, "The Measurement of Drivers ' Mental Workload," *PhD Dissertation, University of Groningen*, 1996.

[21] D. Kahneman, *Attention and effort*, vol. 88, no. 2.New Jersey, U.S.A, Prentice Hall, 1973.

[22] C. D. Wickens, "Multiple resources and performance prediction," *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2. pp. 159–177, 2002.

[23] M. S. Young and N. A. Stanton, "Attention and automation: New perspectives on mental underload and performance," *Theor. Issues Ergon. Sci.*, vol. 3, no. 2, pp. 178–194, 2002.

[24] M. S. Young, K. a Brookhuis, C. D. Wickens, and P. a Hancock, "State of Science: mental workload in ergonomics," *Ergonomics*, vol. 58, no. 1, pp. 1–17, 2015.

[25] D. de Waard, M. Jessurun, F. J. Steyvers, P. T. Raggatt, and K. A. Brookhuis, "Effect of road layout and road environment on driving performance, drivers' physiology and road appreciation," *Ergonomics*, vol. 38, no. 7, pp. 1395–1407, 1995.

[26] S. T. Godley, T. J. Triggs, and B. N. Fildes, "Perceptual lane width, wide perceptual road centre markings and driving speeds.," *Ergonomics*, vol. 47, no. 3, pp. 237–256, 2004.

[27] S. G. Hart, M. F. California, and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Adv. Psychol.*, vol. 52, pp. 139–183, 1988.

[28] F. Rosey, J.-M. Auberlet, O. Moisan, and G. Dupré, "Impact of Narrower Lane Width," *Transportation Research Rec. Journal Transp. Res. Board*, vol. 2138, pp. 112–119, 2009.

[29] C. Dijksterhuis, K. A. Brookhuis, and D. De Waard, "Effects of steering demand on lane keeping behaviour, self-reports, and physiology. A simulator study," *Accid. Anal. Prev.*, vol. 43, no. 3, pp. 1074–1081, 2011.

[30] S. G. Charlton, "Perceptual and attentional effects on drivers' speed selection at

curves," *Accid. Anal. Prev.*, vol. 36, no. 5, pp. 877–884, 2004.

[31] H. Godthelp and W. van Winsum, "Speed Choice and Steering Behavior in Curve Driving," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 38, no. 3, pp. 434–441, 1996.

[32] R. a. Krammes, R. Q. Brackett, M. a Shafer, J. L. Ottesen, I. B. Anderson, K. L. Fink, K. M. Collins, O. J. Pendleton, and C. J. Messer, "Horizontal alignment design consistency for rural two-lane highways. ", *Technical Report, US Department of Transportation*, 1995.

[33] F. Rosey and J. M. Auberlet, "Trajectory variability: Road geometry difficulty indicator," *Safety Science*, vol. 50, no. 9, pp. 1818–1828, 2012.

[34] O. Tsimhoni, H. Yoo, and G. Paul, "Effects of Visual Demand and In-Vehicle Task Complexity on Driving and Task Performance as Assessed by Visual Occlusion," *Technical Report UMTRI - 99 - 37.University of Michigan Transportation Research Institute*, 1999.

[35] M. Vollrath and I. Totzke, "In-Vehicle Communication and Driving : An Attempt to Overcome their Interference," *National Highway Traffic Safety Administration*, 2000. [Online]. Available: http://www-nrd.nhtsa.dot.gov/departments/Human Factors/driver-distraction/Topics043100033.htm.

[36] T. Ben-Bassat and D. Shinar, "Effect of shoulder width, guardrail and roadway geometry on driver perception and behavior," *Accid. Anal. Prev.*, vol. 43, no. 6, pp. 2142–2152, 2011.

[37] C. Ariën, E. M. M. Jongen, K. Brijs, T. Brijs, S. Daniels, and G. Wets, "A simulator study on the impact of traffic calming measures in urban areas on driving behavior and workload," *Accid. Anal. Prev.*, vol. 61, pp. 43–53, 2013.

[38] H. Du, X. Zhao, X. Zhang, Y. Zhang, and J. Rong, "Effects of fatigue on driving performance under different roadway geometries: a simulator study.," *Traffic Inj. Prev.*, vol. 16, no. 5, pp. 468–73, 2014.

[39] J. Son, Y. Lee, and M. H. Kim, "Impact of traffic environment and cognitive workload on older drivers' behavior in simulated driving," *Int. J. Precis. Eng. Manuf.*, vol. 12, no. 1, pp. 135–141, 2011.

[40] B.-S. Liu and Y.-H. Lee, "In-vehicle workload assessment: effects of traffic situations and cellular telephone use.," *J. Safety Res.*, vol. 37, no. 1, pp. 99–105, Jan. 2006.

[41] C. J. D. Patten, A. Kircher, J. Ostlund, L. Nilsson, and O. Svenson, "Driver experience and cognitive workload in different traffic environments.," *Accid. Anal. Prev.*, vol. 38, no. 5, pp. 887–94, Sep. 2006.

[42] A. Stinchcombe and S. Gagnon, "Driving in dangerous territory: Complexity and road-characteristics influence attentional demand," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 13, no. 6, pp. 388–396, 2010.

[43] D. L. Strayer, F. a Drews, and W. a Johnston, "Cell phone-induced failures of visual attention during simulated driving.," *J. Exp. Psychol. Appl.*, vol. 9, no. 1, pp. 23–32, 2003.

[44] X. Hao, Z. Wang, F. Yang, Y. Wang, and Y. Guo, "The Effect of Traffic on Situation Awareness and Mental Workload：Simulator-Based Study," *Engin. Psychol. Cog. Ergon.*, pp. 288–296, 2007.

[45] K. A. Brookhuis and D. de Waard, "Monitoring drivers' mental workload in driving simulators using physiological measures," *Accid. Anal. Prev.*, vol. 42, pp. 898–903, 2010.

[46] T. C. Lansdown, N. Brook-Carter, and T. Kersloot, "Distraction from multiple in-vehicle secondary tasks: vehicle performance and mental workload implications.," *Ergonomics*, vol. 47, no. 1, pp. 91–104, 2004.

[47] M. P. Reed and P. A. Green, "Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task," *Ergonomics*, vol. 42, no. 8, pp. 1015–1037, 1999.

[48] A.-S. Wikman, T. Nieminen, and H. Summala, "Driving experience and time-sharing during in-car tasks on roads of different width," *Ergonomics*, vol. 41, no. 3, pp. 358–372, 1998.

[49] D. E. Crundall and G. Underwood, "Effects of Experience and Processing Demands on Visual Information Acquisition in Drivers," *Ergonomics*, vol. 41, no. 4, pp. 448–458, 1998.

[50] C. J. D. Patten, A. Kircher, J. Östlund, L. Nilsson, and O. Svenson, "Driver experience and cognitive workload in different traffic environments," *Accid. Anal. Prev.*, vol. 38, no. 5, pp. 887–894, 2006.

[51] Y. Yang, "The Effects of Increased Workload on Driving Performance and Visual Behaviour," *PhD Dissertation, University of Southampton,* 2011.

[52] O. Nakayama, T. Futami, T. Nakamura, and E. R. Boer, "Development of a Steering Entropy Method for Evaluating Driver Workload," *Proc. JSAE Annu. Congr.*, no. 724, pp. 5–8, 1999.

[53] J. Paxion, E. Galy, and C. Berthelon, "Mental Workload and Driving," *Front. Psychol.*, vol. 5, no. December, pp. 1–11, Dec. 2014.

[54] K. A. Brookhuis, G. de Vries, and D. de Waard, "The effects of mobile telephoning on driving performance," *Accid. Anal. Prev.*, vol. 23, no. 4, pp. 309–316, 1991.

[55] R.D. O'Donnel and F.T. Eggemeier, *, "Workload assessment methodology," in Handbook of Perception and Human Performance, Vol II, Cognitive Processes and Performance,*. Wiley, New York, 1986.

[56] G. Mulder, "The concepts and measurement of mental effort," in *Volume 31 of the series NATO ASI Series*, 1986, pp. 175–198.

[57] L. J. M. Mulder, "Measurement and analysis methods of heart rate and respiration for use in applied environments," *Biol. Psychol.*, vol. 34, no. 2–3, pp. 205–236, 1992.

[58] R. H. Mulder B, Veldman H, Van der Veen F, Van Roon A and M. B. Schachinger H, "On the effects of mental task performance on heart rate, blood pressure and its variability measures," *Blood Press. Hear. Rate Var.*, pp. 153–166, 1992.

[59] B. Mulder, H. Rusthoven, M. Kuperus, M. de Rivecourt, and D. de Waard, "Short-

term heart rate measures as indices of momentary changes in invested mental effort," *Hum. Factors Issues Complex Syst. Perform.*, pp. 1–16, 2007.

[60] A. Stuiver, K. A. Brookhuis, D. de Waard, and B. Mulder, "Short-term cardiovascular measures for driver support: Increasing Sensitivity for Detecting Changes in Mental Workload," *Int. J. Psychophysiol.*, vol. 92, pp. 35–41, 2014.

[61] N. Ochiai, E. Yasuda, K. Murata, and Y. Ueno, "Development of Simplified Appraisal Method of Fatigue on Sitting for Extended Periods by the Data of Finger Plethysmogram," *Japanese J. Ergon.*, vol. 40, no. 5, pp. 254–263, 2004.

[62] Tokyo Web, "Electrocardiogram check by car handle that predict heart attack," *Tokyo Shinbun*, Tokyo, 2013.

[63] S. Miller, "Workload Measures," *Technical Report, National Advanced Driving Simulator IOWA City, United States,* 2001.

[64] F.R.H. Zijlstra, "Efficiency in Work Behavior (Doctoral Dissertation)," *Technical University Delft, The Netherlands.*, 1993.

[65] G. B. Reid, S. S. Potter, and J. R. Bressler, "Subjective Workload Assessment Technique (SWAT): A scaling procedure for measuring mental workload.," in *Human Mental Workload*, P. A. Hancock & N. Meshkati (Eds), Ed. Amsterdam: Elsevier, 1988, pp. 185–218.

[66] S. Rubio, E. Díaz, J. Martín, and J. M. Puente, "Evaluation of Subjective Mental Workload : A Comparison of SWAT , NASA-TLX , and Workload Profile Methods," vol. 53, no. 1, pp. 61–86, 2004.

[67] O. S. Tiejun MIAO, Toshiyuki SHIMIZU, "The Use of Chaotic Dynamics in Finger Photoplethysmography to Assess Driver Mental Workload," in *The Society of Automotive Engineers of Japan*, 2003, no. 18–3, pp. 18–3, 1--3.

[68] K. Suzuki and Y. Okada, "Evaluation of Driver's Mental Workload in Terms of the Fluctuation of Finger Pulse," *Trans. JSME*, vol. 74, no. 743, pp. 85–94, 2008.

[69] D. M. A. Gronwall, "Paced Auditory Task: A Measure of Recovery From COnsussion," *Perceptual Mot. Ski.*, vol. 44, pp. 367–373, 1977.

[70] S. Enokida, K. Kotani, S. Suzuki, and T. Asao, "Assessing Mental Workload of In-Vehicle Information," in *Human Interface and the Management of Information. Information and Interaction Design*, vol. 8016, 2013, pp. 584–593.

[71] K. Suzuki and Y. Okada, "Evaluation of Driver's Mental Workload in terms of the fluctuation of finger pulse (in Japanese)," in *The Japan Society of Mechanical Engineers*, 2007, pp. 293–296.

[72] H. Alm and L.Nilsson, "The Effects of A Mobile Telephone Task on Driver Behaviour in a Car Following Situation," *Accid. Anal. Prev.*, vol. 27, no. 5, pp. 707–715, 1995.

[73] G. Jahn, A. Oehme, J. F. Krems, and C. Gelau, "Peripheral Detection as a Workload Measure in Driving: Effects of Traffic Complexity and Route Guidance System Use in a Driving Study," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 8, no. 3, pp. 255–275, 2005.

[74] D. L. Strayer, J. M. Cooper, J. Turrill, J. Coleman, and N. Medeiros-, "Measuring

Cognitive Distraction in the Automobile," *Technical Report, Foundation for Traffic Safety, Washinton DC,* June, 2013.

[75] G. Marquart and J. de Winter, "Workload assessment for mental arithmetic tasks using the task-evoked pupillary response," *Peer J. Comput. Sci.*, vol. 1, p. e16, 2015.

[76] F. Platten, "Analysis of Mental Workload and Operating Behavior in Secondary Tasks while Driving," *PhD Dissertation, Shemnitz University of Technology,* 2012.

[77] M. Pagani, F. Lombardi, S. Guzzetti, O. Rimoldi, R. Furlan, P. Pizzinelli, G. Sandrone, G. Malfatto, S. Dell'Orto, and E. Piccaluga, "Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympathovagal interaction in man and conscious dog.," *Circ. Res.*, vol. 59, no. 2, pp. 178–193, 1986.

[78] A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti, "Cardiovascular neural regulation explored in the frequency domain.," *Circulation*, vol. 84, no. 2, pp. 482–492, 1991.

[79] J. P. A. Delaney AND D. A. Brodie, "Effects of Short-Term Psychological Stress on the Time and Frequency Domains of Heart-Rate Variability," *Percept. Mot. Skills*, vol. 91, pp. 515–524, 2000.

[80] R. Orsila, M. Virtanen, T. Luukkaala, M. Tarvainen, P. Karjalainen, J. Viik, M. Savinainen, and C.-H. Nygård, "Perceived mental stress and reactions in heart rate variability--a pilot study among employees of an electronics company.," *Int. J. Occup. Saf. Ergon.*, vol. 14, no. 3, pp. 275–283, 2008.

[81] C. D. B. Luft, E. Takase, and D. Darby, "Heart rate variability and cognitive function: Effects of physical effort," *Biol. Psychol.*, vol. 82, no. 2, pp. 196–201, 2009.

[82] B. Cinaz, R. La Marca, and B. Arnrich, "Monitoring of mental workload levels," in *IADIS e-Health Conference*, 2010, pp. 189–193.

[83] W. J. Horrey, M. F. Lesch, and a. Garabet, "Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions," *J. Safety Res.*, vol. 40, no. 1, pp. 7–12, 2009.

[84] N. Hjortskov, D. Rissén, A. K. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work.," *Eur. J. Appl. Physiol.*, vol. 92, no. 1–2, pp. 84–9, Jun. 2004.

[85] R. W. Backs, J. K. Lenneman, J. M. Wetzel, and P. Green, "Cardiac measures of driver workload during simulated driving with and without visual occlusion.," *Hum. Factors*, vol. 45, no. 4, pp. 525–538, 2015.

[86] J.A. Veltman and A.W.K Gaillard, "Physiological workload reactions to increasing levels of task," *Ergonomics,* pp. 656–670, 1998.

[87] C. Dijksterhuis, "Monitoring Driver ' s Mental Workload for User Adaptive Aid," *PhD Dissertation, University of Groningen*, 2014.

[88] W. B. Verwey, "On-line driver workload estimation. Effects of road situation and age on secondary task measures," *Ergonomics*, Vol. 43, No. 2, pp.187-209, 2000.